

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2002, 12:38:07 : Search time 32.65 seconds
(without alignments)
1126.047 Million cell updates/sec

Title: US-09-884-948-1
Perfect score: 1811
Sequence: 1 DSDRVTTPAEPLDRMPDPY.....ITFIPKSWNTAPDKVKQGP 331

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802.*

1:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1811	100.0	331	20	AAV33662
2	1811	100.0	331	20	AAV33665
3	1811	100.0	331	20	AAW67770
4	1811	100.0	331	22	AAAB81161
5	1811	100.0	332	20	AAW67771
6	1811	100.0	332	21	AAAB12809
7	1811	100.0	346	15	AAAR49048
8	1811	100.0	406	13	AAAR22651
9	1811	100.0	407	22	AAAB97831
10	1811	100.0	407	22	AAAB91166
11	1811	100.0	408	15	AAAR49049

Transglutaminase (Transglutaminase r S. cinnamomeum IFO Transglutaminase Streptovorticilliu Transglutaminase N Peptide fragment o Peptide fragment o Peptide fragment o Peptide fragment o Peptide fragment o Drosophila melanog Drosophila melanog AAV3A capsid prote Human cancer assoc Human metastasis-a AAV3B capsid prote Adeno-associated v AAV6 capsid protei Peptide fragment o Arabidopsis thalia Arabidopsis thalia Arabidopsis thalia Peptide fragment o 1-Caldesmon, Gall Adeno-associated v Adeno-associated v AAV2 capsid protei Peptide fragment o C albicans apoptos Drosophila melanog

Transglutaminase (Transglutaminase r S. cinnamomeum IFO Transglutaminase Streptovorticilliu Transglutaminase N Peptide fragment o Peptide fragment o Peptide fragment o Peptide fragment o Peptide fragment o Drosophila melanog Drosophila melanog AAV3A capsid prote Human cancer assoc Human metastasis-a AAV3B capsid prote Adeno-associated v AAV6 capsid protei Peptide fragment o Arabidopsis thalia Arabidopsis thalia Arabidopsis thalia Peptide fragment o 1-Caldesmon, Gall Adeno-associated v Adeno-associated v AAV2 capsid protei Peptide fragment o C albicans apoptos Drosophila melanog

Transglutaminase (Transglutaminase r S. cinnamomeum IFO Transglutaminase Streptovorticilliu Transglutaminase N Peptide fragment o Peptide fragment o Peptide fragment o Peptide fragment o Peptide fragment o Drosophila melanog Drosophila melanog AAV3A capsid prote Human cancer assoc Human metastasis-a AAV3B capsid prote Adeno-associated v AAV6 capsid protei Peptide fragment o Arabidopsis thalia Arabidopsis thalia Arabidopsis thalia Peptide fragment o 1-Caldesmon, Gall Adeno-associated v Adeno-associated v AAV2 capsid protei Peptide fragment o C albicans apoptos Drosophila melanog

ALIGNMENTS

RESULT 1

AAV33662

ID AAV33662 standard; Protein; 331 AA.

XX

AC AAV33662;

DT DT

XX 07-JAN-2000 (first entry)

XX Streptovorticillium S-8112 transglutaminase protein fragment.

XX Transglutaminase; food industry; pharmaceutical industry; texture; cosmetic industry; proteinaceous material; gel strength; viscosity; breaking strength; elasticity; taste.

XX Streptovorticillium sp.

XX DE19814860-A1.

XX 07-OCT-1999.

XX 02-APR-1998; 98DE-1014860.

XX 02-APR-1998; 98DE-1014860.

XX (FUCH/) FUCHSBAUER H.

XX Fuchsbauser H, Pasternack R, Dorsch S, Otterbach J, Robenek I; Mainusch M, Dauscher C;

XX WPI: 1999-552288/47.

XX N-PSDB; AA223653.

XX Bacterial transglutaminase polypeptides useful for polymerizing proteins, e.g. to modify the properties of food, pharmaceutical or

PT cosmetic products -
 XX Claim 1; Page 23-24; 44pp; German.
 PS
 CC This invention describes a novel bacterial transglutaminase polypeptide.
 CC (1) can be used, e.g. in the food, pharmaceutical and cosmetic
 CC industries, to polymerize proteinaceous materials in order to improve
 CC their properties, e.g. texture, gel strength, breaking strength,
 CC viscosity, elasticity or taste. (1) can also be used to immobilize
 CC enzymes and antibodies. This sequence represents a transglutaminase
 CC isolated from Streptococcus sp. S-8112.
 XX Sequence 331 AA;
 SQ

Query Match 100.0%; Score 1811; DB 20; Length 331;
 Best Local Similarity 100.0%; Pred. No. 2.3e-153;
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSDRVTTPAEPDLPDPYPSYGRATVNNYIRKQVYSHRDKRQKQMTTEQREWL 60
 DB 1 DSDRVTTPAEPDLPDPYPSYGRATVNNYIRKQVYSHRDKRQKQMTTEQREWL 60
 QY 61 SYGCVGTWVNSGYPTNRLAFASDFEDRFKNEKNGRPSGTRAEFEGRVAKESFDEE 120
 DB 61 SYGCVGTWVNSGYPTNRLAFASDFEDRFKNEKNGRPSGTRAEFEGRVAKESFDEE 120
 QY 121 KGFQREAVASVMNRALENANDESAVLDNLKELANGNDALRNEDARSFYSAIRNTPSF 180
 DB 121 KGFQREAVASVMNRALENANDESAVLDNLKELANGNDALRNEDARSFYSAIRNTPSF 180
 QY 181 KERNNGNDPSPRMKAVIYKHFWSGQDRSSADKRYGDPDAFRPAGTGLVDMSRDRI 240
 DB 181 KERNNGNDPSPRMKAVIYKHFWSGQDRSSADKRYGDPDAFRPAGTGLVDMSRDRI 240
 QY 241 PRSPTSGEGFVNFYGFQAGTQADAKTVWTHGNHYHAPNGSLGAMHYVESKFRNWE 300
 DB 241 PRSPTSGEGFVNFYGFQAGTQADAKTVWTHGNHYHAPNGSLGAMHYVESKFRNWE 300
 QY 301 GYSDFDRGAYVITFIPKSNMTAPDKVKQGW 331
 DB 301 GYSDFDRGAYVITFIPKSNMTAPDKVKQGW 331

RESULT 2
 ID AAY33665
 AC AAY33665;
 DT 07-JAN-2000 (first entry)
 DE Streptococcus sp. molarase transglutaminase protein fragment.
 XX Transglutaminase; food industry; pharmaceutical industry; texture;
 KW cosmetic industry; proteinaceous material; gel strength; viscosity;
 KW breaking strength; elasticity; taste.
 XX Streptococcus sp. molarase.
 OS DE19814860-A1.
 PN 07-OCT-1999.
 PD 02-APR-1998; 98DE-1014860.
 PF 02-APR-1998; 98DE-1014860.
 PR (FUCHS) FUCHSBAUER H.
 PA Fuchsbauer H., Pasternack R., Dorsch S., Otterbach J., Robenek I;
 PI Mainusch M., Dauscher C;
 XX

DR WPI; 1999-552288/47.
 XX Bacterial transglutaminase polypeptides useful for polymerizing
 PT proteins, e.g. to modify the properties of food, pharmaceutical or
 PT cosmetic products -
 XX Claim 12; Page 25-26; 44pp; German.
 PS
 CC This invention describes a novel bacterial transglutaminase polypeptide.
 CC (1) can be used, e.g. in the food, pharmaceutical and cosmetic
 CC industries, to polymerize proteinaceous materials in order to improve
 CC their properties, e.g. texture, gel strength, breaking strength,
 CC viscosity, elasticity or taste. (1) can also be used to immobilize
 CC enzymes and antibodies. This sequence represents a transglutaminase
 CC protein fragment isolated from Streptococcus sp. molarase.
 XX Sequence 331 AA;
 SQ

Query Match 100.0%; Score 1811; DB 20; Length 331;
 Best Local Similarity 100.0%; Pred. No. 2.3e-153;
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSDRVTTPAEPDLPDPYPSYGRATVNNYIRKQVYSHRDKRQKQMTTEQREWL 60
 DB 1 DSDRVTTPAEPDLPDPYPSYGRATVNNYIRKQVYSHRDKRQKQMTTEQREWL 60
 QY 61 SYGCVGTWVNSGYPTNRLAFASDFEDRFKNEKNGRPSGTRAEFEGRVAKESFDEE 120
 DB 61 SYGCVGTWVNSGYPTNRLAFASDFEDRFKNEKNGRPSGTRAEFEGRVAKESFDEE 120
 QY 121 KGFQREAVASVMNRALENANDESAVLDNLKELANGNDALRNEDARSFYSAIRNTPSF 180
 DB 121 KGFQREAVASVMNRALENANDESAVLDNLKELANGNDALRNEDARSFYSAIRNTPSF 180
 QY 181 KERNNGNDPSPRMKAVIYKHFWSGQDRSSADKRYGDPDAFRPAGTGLVDMSRDRI 240
 DB 181 KERNNGNDPSPRMKAVIYKHFWSGQDRSSADKRYGDPDAFRPAGTGLVDMSRDRI 240
 QY 241 PRSPTSGEGFVNFYGFQAGTQADAKTVWTHGNHYHAPNGSLGAMHYVESKFRNWE 300
 DB 241 PRSPTSGEGFVNFYGFQAGTQADAKTVWTHGNHYHAPNGSLGAMHYVESKFRNWE 300
 QY 301 GYSDFDRGAYVITFIPKSNMTAPDKVKQGW 331
 DB 301 GYSDFDRGAYVITFIPKSNMTAPDKVKQGW 331

RESULT 3
 ID AAW67770
 AC AAW67770;
 DT 01-APR-1999 (first entry)
 DE A transglutaminase enzyme sequence.
 XX Transglutaminase; microbial; gelled food; jelly; yogurt; cheese;
 KW cosmetic; meat quality; microcapsule production; high thermal stability;
 KW carrier; immobilised enzyme.
 XX Streptococcus sp.
 OS EP889133-A2.
 PN 07-JAN-1999.
 PD 02-JUL-1998; 98EP-0112315.
 PF 04-JUL-1997; 97JP-0180010.
 PR (AJIN) AJINOMOTO CO INC.
 PA

```

XX MIWA T, Nakamura N, Seguro K, Yokoyama K;
PI WPI; 1999-062664/06.
XX N-PSDB; AAW81507.
XX New microbial transglutaminase with N-terminal aspartic acid deleted
PT - allowing high level recombinant production without added
PT methionine in E. coli, useful in production of gelled foods,
PT cosmetics etc.
XX
XX Claim 1; Page 12-14; 56pp; English.
XX
XX The present sequence represents a transglutaminase of Streptovorticillum
CC sp.. The specification describes a new microbial transglutaminase that
CC has the N-terminal aspartic acid of the present transglutaminase
CC deleted. Eliminating the N-terminal Asp from microbial transglutaminase
CC allows efficient removal of the terminal Methionine residue added when
CC the protein is expressed in Escherichia coli. The E. coli methionine
CC aminopeptidase acts well on Met-Ser but only poorly on Met-Asp, so
CC problems of antigenicity associated with Met-terminated proteins are
CC avoided. Recombinant transglutaminase is used to produce gelled foods
CC (Jellies, yogurt and cheeses) or cosmetics, to improve the quality of
CC meat, in the production of materials for microcapsules of high thermal
CC stability and as a carrier for immobilised enzymes.
XX
XX Sequence 331 AA;
SQ
Query Match 100.0%; Score 1811; DB 20; Length 331;
Best Local Similarity 100.0%; Pred. No. 2.3e-153;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DSDDRVTPPAELDRMPDPYRPSYGRAETVYNNYIRKWOQVYSHRDGRKQKQMTTEQREWL 60
DB 1 dssdrvtppepldrmpdpypysygraetvnnvirkwqvyshrdgrkqmtteeqrewl 60
QY 61 SYGCVGTWVNSGOYPTNRLAFASDFEDRFKNELKNGRPSGTRAEFGRVAKESFDEE 120
DB 61 sygcvgvtwnsgyptnrlafasdfedrknlnkngprsgtraeefgrvakesfdee 120
QY 121 KGFQREAVASVMNRALENAHDESAYLDNLKELANGNDALRNEDARSFYSLRNTPSF 180
DB 121 kgfqrarevasvmnralenahdesayldnlkkelangndalrnedarspfysalrntpsf 180
QY 181 KERNGNHDPSPRMKAVIYSKHFWSGQDRSSADKRYGDPDAFRPAGTGLVDMSRDRI 240
DB 181 kerngnhdpssrmkaviyskhfwsqgdrssadkrkygdpdafrpagtglvdmrdrni 240
QY 241 PRSPTSPGEGFVNFYDYGWFGAQTADADKTVTHGNHYHAPNGSLGAMHYVESKFRNWE 300
DB 241 prsptspgegfvnfydygwfgaqteadadktvthgnhyhapngslgamhyveskfrnwse 300
QY 301 GYSDFDRGAYVITFIPKSNWNTAPDKVKQGW 331
DB 301 gysdfdrgayvitfipkswnatapdkvkqgwp 331
RESULT 4
AAB81161
ID AAB81161 standard; protein: 331 AA.
XX
XX AAB81161;
XX
XX 13-JUL-2001 (first entry)
XX
XX Transglutaminase related protein SEQ ID 5.
XX
XX Coryneform bacteria; transglutaminase; food processing.
XX
XX Streptovorticillum mobaraense.
XX
XX WO200123591-A1.
PN

```

```

XX 05-APR-2001.
XX
XX 29-SEP-2000; 2000WO-JP06780.
XX
XX 30-SEP-1999; 99JP-0280098.
XX
XX 28-JUN-2000; 2000JP-0194043.
XX
XX (AJIN ) AJINOMOTO CO INC.
XX
XX Kikuchi Y, Date M, Umezawa Y, Yokoyama K, Matsui H;
XX
XX WPI; 2001-266172/27.
XX
XX Efficient secretory production of foreign proteins e.g.
XX transglutaminase employing transformant coryneform bacterium, simply on
XX industrial scale with direct recovery for use in food processing and
XX pharmaceutical industry
XX
XX Claim 22; Page 79-81; 151pp; Japanese.
XX
XX This invention relates to a process for the production of a foreign
XX secretory protein through the construction of a recombinant coryneform
XX bacterium. The coryneform bacterium is transformed with an expression
XX construct in which DNA encoding a target foreign protein pro-structure is
XX ligated to the downstream region of DNA encoding the signal peptide
XX domain of a coryneform bacterial protein. Following transformation with
XX the vector, the bacterium is cultured, and the pro-peptide cleaved from
XX the expressed protein. Transglutaminases produced using this process are
XX useful in the food processing and pharmaceutical industries. The present
XX sequence represents a transglutaminase related protein, which can be
XX used in the method of the invention.
XX
XX Sequence 331 AA;
SQ
Query Match 100.0%; Score 1811; DB 22; Length 331;
Best Local Similarity 100.0%; Pred. No. 2.3e-153;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DSDDRVTPPAELDRMPDPYRPSYGRAETVYNNYIRKWOQVYSHRDGRKQKQMTTEQREWL 60
DB 1 dssdrvtppepldrmpdpypysygraetvnnvirkwqvyshrdgrkqmtteeqrewl 60
QY 61 SYGCVGTWVNSGOYPTNRLAFASDFEDRFKNELKNGRPSGTRAEFGRVAKESFDEE 120
DB 61 sygcvgvtwnsgyptnrlafasdfedrknlnkngprsgtraeefgrvakesfdee 120
QY 121 KGFQREAVASVMNRALENAHDESAYLDNLKELANGNDALRNEDARSFYSLRNTPSF 180
DB 121 kgfqrarevasvmnralenahdesayldnlkkelangndalrnedarspfysalrntpsf 180
QY 181 KERNGNHDPSPRMKAVIYSKHFWSGQDRSSADKRYGDPDAFRPAGTGLVDMSRDRI 240
DB 181 kerngnhdpssrmkaviyskhfwsqgdrssadkrkygdpdafrpagtglvdmrdrni 240
QY 241 PRSPTSPGEGFVNFYDYGWFGAQTADADKTVTHGNHYHAPNGSLGAMHYVESKFRNWE 300
DB 241 prsptspgegfvnfydygwfgaqteadadktvthgnhyhapngslgamhyveskfrnwse 300
QY 301 GYSDFDRGAYVITFIPKSNWNTAPDKVKQGW 331
DB 301 gysdfdrgayvitfipkswnatapdkvkqgwp 331
RESULT 5
AAW67771
ID AAW67771 standard; Protein: 332 AA.
XX
XX AAW67771;
XX
XX 01-APR-1999 (first entry)
XX
XX

```

DE Protein encoded by high expression transglutaminase gene.
 XX Transglutaminase; microbial; gelled food; jelly; yogurt; cheese;
 KW cosmetic; meat quality; microcapsule production; high thermal stability;
 KW carrier; immobilised enzyme.
 XX
 XX Synthetic.
 OS Streptovorticillum sp.
 XX
 PN EP889133-A2.
 XX
 XX 07-JAN-1999.
 XX
 XX 02-JUL-1998; 98EP-0112315.
 XX
 XX 04-JUL-1997; 97JP-0180010.
 XX
 XX (AJIN) AJINOMOTO CO INC.
 PA
 PI Miwa T, Nakamura N, Seguro K, Yokoyama K;
 DR WPI: 1999-062664/06.
 DR N-PSDB; AAV81508.
 XX
 PT New microbial transglutaminase with N-terminal aspartic acid deleted
 PT - allowing high level recombinant production without added
 PT methionine in E. coli, useful in production of gelled foods,
 PT cosmetics etc.
 XX
 PS Example 1; Page 18-23; 56pp; English.
 XX
 XX The present sequence is encoded by the high expression transglutaminase
 CC gene present in plasmid pTGMG-02. The gene is derived from
 CC Streptovorticillum sp., and is codon altered, using oligonucleotides
 CC AAV81521-60, for expression in Escherichia coli. The specification
 CC describes a new microbial transglutaminase that has the N-terminal
 CC aspartic acid of transglutaminase deleted. Eliminating the N-terminal
 CC Asp from microbial transglutaminase allows efficient removal of the
 CC terminal Met residue added when the protein is expressed in E. coli.
 CC The E. coli methionine aminopeptidase acts well on Met-Ser but only
 CC poorly on Met-Asp, so problems of antigenicity associated with
 CC Met-terminated proteins are avoided. Recombinant transglutaminase is
 CC used to produce gelled foods (jellies, yogurt and cheeses) or cosmetics,
 CC to improve the quality of meat, in the production of materials for
 CC microcapsules of high thermal stability and as a carrier for immobilised
 CC enzymes.
 XX
 XX Sequence 332 AA;
 SQ
 Query Match 100.0%; Score 1811; DB 20; Length 332;
 Best Local Similarity 100.0%; Pred. No. 2.3e-153;
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DSDRVTTPAEPLDRMPDPYRPSYGRAETVNNYIRKQOQVYSHRDGRKQOQTEQREWL 60
 DB 2 dsddrvtpaepldrmpdpypsygraetvnnyrkwwqvyshrdgrkqgteeqrewl 61
 QY 61 SYGCVGVTVWNSGQYPTNRLAFASFDDEFRKLNKNGRPSRGETRAEFGRVAKESFDEE 120
 DB 62 sygcgvgtwvnsqyptnrlafasfdedfrklnkngprsgetrtaefgrvakesfdee 121
 QY 121 KGFQREAVASVNNRALENADHESAYLDNLKKELANGNDALRNEDARSPFYSALRNTPSF 180
 DB 122 kgfqrarevasvnnralenadhesayldnlkkelangndalrnedarspfysalrntpsf 181
 QY 181 KERNGNHFDSRMKAVIYKHFWSGDRSSADKRYGPDPAFRPAGTGLVDMRSRDN 240
 DB 182 kerngnhdpfsmkavlyskhfwsqgdrssadkrygpdptrpgtgtlvdmrsdrni 241
 QY 241 PRSTPSGEGVNFDFYGFCAQTEADAKTWTGHNYHAPNGSLGAMHVYESKPRNWE 300
 DB 242 prstpsgegfvnfdygfcaqteadaktwtghnyhapngslgamhvyeskfwnwe 301

QY 301 GYSDFDRGAYVITFIKPSWNTAPDKVKQGW 331
 DB 302 gysdfdrgayvitfipkswntapdkvkqgwp 332
 RESULT 6
 AAB12809
 ID AAB12809 standard; Protein; 332 AA.
 XX
 AC AAB12809;
 XX
 DT 24-NOV-2000 (first entry)
 DE Transglutaminase protein sequence SEQ ID NO:1.
 XX
 KW Transglutaminase; gelled food; jelly; yoghurt; gelled cosmetic;
 KW cheese.
 XX
 OS Unidentified.
 XX
 PN WO200040706-A1.
 XX
 PD 13-JUL-2000.
 XX
 PF 24-DEC-1999; 99WO-JP07250.
 XX
 PR 28-DEC-1998; 98JP-0373131.
 XX
 PA (AJIN) AJINOMOTO CO INC.
 XX
 XX Yokoyama K, Ono K, Ejima D;
 PI
 XX WPI: 2000-475826/41.
 DR N-PSDB; AAA73025.
 XX
 XX Production of active transglutaminase from denatured enzyme by
 PT two-stage refolding process for industrial production of active enzyme
 PT for use in food production -
 XX
 PS Disclosure; Page 48-50; 74pp; Japanese.
 XX
 XX The present invention describes a method for producing active
 CC transglutaminase from denatured enzyme. The method comprises: (i) forming
 CC an intermediate structure of the enzyme having transglutaminase activity
 CC under acidic conditions in an aqueous medium; and (ii) forming a higher-
 CC level structure of the enzyme having transglutaminase activity under
 CC neutral conditions in an aqueous medium. The method can be used for
 CC industrial production of active transglutaminase from denatured material
 CC (such as recombinant transglutaminase) which can be used in the food
 CC industry for the production of gelled foods such as jellies, yoghurts
 CC and cheeses, and for the production of gelled cosmetics. The present
 CC sequence represents a transglutaminase which is used in the
 CC exemplification from the present invention.
 XX
 SQ Sequence 332 AA;
 Query Match 100.0%; Score 1811; DB 21; Length 332;
 Best Local Similarity 100.0%; Pred. No. 2.3e-153;
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DSDRVTTPAEPLDRMPDPYRPSYGRAETVNNYIRKQOQVYSHRDGRKQOQTEQREWL 60
 DB 2 dsddrvtpaepldrmpdpypsygraetvnnyrkwwqvyshrdgrkqgteeqrewl 61
 QY 61 SYGCVGVTVWNSGQYPTNRLAFASFDDEFRKLNKNGRPSRGETRAEFGRVAKESFDEE 120
 DB 62 sygcgvgtwvnsqyptnrlafasfdedfrklnkngprsgetrtaefgrvakesfdee 121
 QY 121 KGFQREAVASVNNRALENADHESAYLDNLKKELANGNDALRNEDARSPFYSALRNTPSF 180
 DB 122 kgfqrarevasvnnralenadhesayldnlkkelangndalrnedarspfysalrntpsf 181

```

QY 181 KERNGNHDPSPRMKAVIYKSHFWGQDRSSADKRYKGGPDPAFRPAPGTGLVDMSRDRI 240
    |||||||
Db 182 kerngnhdpSRmkavIyskhfwsqgdrssadkrkygpdafpapgTglvdmSRdrni 241
    |||||||
QY 241 PRSPTSPGEGFVNFdyGwfgaQTEADADKTVTHGNHYHAPNGSLGAMHVYESKFRNWSE 300
    |||||||
Db 242 prsptspgEgfvnfdygwgfgaQTEADADKTVTHGNHYHAPNGSLGAMHVYESKFRNWSE 301
    |||||||
QY 301 GYSDFDRGAYVITFIPKSWNTAPDKVKQGW 331
    |||||||
Db 302 gysdfdrGAYVITfipkswntapdkvkqgw 332
    |||||||

RESULT 7
AAR49048
ID AAR49048 standard; Protein; 346 AA.
XX
AC AAR49048;
XX
XX 20-SEP-1994 (first entry)
XX
XX Bacterial transglutaminase.
XX
KW Bacterial transglutaminase; BTG; expression;
KW active; inactive; inclusion body.
XX
PN JP06030771-A.
XX
PD 08-FEB-1994.
XX
XX 14-JUL-1992; 92JP-0187038.
XX
XX 14-JUL-1992; 92JP-0187038.
XX
PA (AJIN ) AJINOMOTO KK.
XX
DR WPI: 1994-079294/10.
DR N-PSDB; AAQ55983.
XX
XX Prepn. of bacterial trans-glutaminase in large amts. - by
PT expression of fusion protein in E. coli bacterial
PT trans-glutaminase
XX
PS Claim 1; Page 7-8; 13pp; Japanese.
XX
XX A fused protein contains amino acids 16-346 of BTG (AAQ55983)
CC and a hydrophilic peptide at the amino terminal.
CC Expression of DNA encoding this protein in E. coli allow large
CC scale prodn. of BTG. An active BTG can be prepd. from the
CC inactive fused protein inclusion body.
XX
XX Sequence 346 AA;

Query Match 100.0%; Score 1811; DB 15; Length 346;
Best Local Similarity 100.0%; Pred. No. 2.4e-153;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSDDRVTTPAEPLDRMDPPYRPSYGRAETVNNYKRWQVYSHRGRKQWTEQREWL 60
    |||||||
Db 16 dsddrvtpaEPLdrmpdpYrpsYgraetvnnYirkwqVysHrgrkqWteeqrewl 75
    |||||||
QY 61 SYGCVGVTVWNSGQYPNRLAFASDFDRFKNELKNGRPSRGETRAEFEGRVAKESFDEE 120
    |||||||
Db 76 sygcvgvTVwnsgqypnrlafasdfdrfknelkngRpsrgetraefegrvakesfdee 135
    |||||||

QY 121 KGFORAREVASVNNRLENADHDSAYLDNLKKELANGDNLARNEDARSPPYSALRNTPSF 180
    |||||||
Db 136 kgfdrarevasvnnrAlenahdesayldnlkkelangdnLARNedarsppysalrntpsf 195
    |||||||
QY 181 KERNGNHDPSPRMKAVIYKSHFWGQDRSSADKRYKGGPDPAFRPAPGTGLVDMSRDRI 240
    |||||||

```

```

Db 196 kerngnhdpSRmkavIyskhfwsqgdrssadkrkygpdafpapgTglvdmSRdrni 255
    |||||||
QY 241 PRSPTSPGEGFVNFdyGwfgaQTEADADKTVTHGNHYHAPNGSLGAMHVYESKFRNWSE 300
    |||||||
Db 256 prsptspgEgfvnfdygwgfgaQTEADADKTVTHGNHYHAPNGSLGAMHVYESKFRNWSE 315
    |||||||
QY 301 GYSDFDRGAYVITFIPKSWNTAPDKVKQGW 331
    |||||||
Db 316 gysdfdrGAYVITfipkswntapdkvkqgw 346
    |||||||

RESULT 8
AAR22651
ID AAR22651 standard; Protein; 406 AA.
XX
AC AAR22651;
XX
XX 09-OCT-1992 (first entry)
XX
XX Transglutaminase.
XX
KW BTG; acyl rearrangement; deamination.
XX
OS Actinomycetes and Streptovorticillium.
XX
FH Key Location/Qualifiers
FT Peptide 1..75
FT /label= sig_peptide
FT Protein 76..406
FT /label= mat_BTG
XX
XX EP481504-A.
XX
XX 22-APR-1992.
XX
XX 18-OCT-1991; 91EP-0117813.
XX
XX 19-OCT-1990; 90JP-0282566.
XX
XX (AJIN ) AJINOMOTO CO INC.
XX (AMAN ) AMANO PHARM KK.
XX
XX Ando K, Arafuka S, Koikeda S, Matsui H, Takagi H, Washizu K;
XX WPI: 1992-133808/17.
XX N-PSDB; AAQ24197 and AAQ24201.
XX
XX DNA fragment encoding trans:glutaminase - is inserted into
PT vector, e.g. pNJI053-BTG, for protein expression
XX
XX Disclosure; Page 42 and 44; 55 pp; English.
XX
XX The mature transglutaminase enzyme (BTG) can be derived from two
CC different species, Streptovorticillium and Actinomycetes. It
CC catalyses an acyl rearrangement reaction of a gamma-carboxamide gp.
CC of glutamine. It introduces intra- or intermolecular formation of
CC epsilon-(gamma-Gln)-Lys cross-linking when an epsilon-amino gp. of
CC a Lys residue acts as an acyl receptor. When water acts as an acyl
CC acceptor the enzyme accelerates the conversion of Gln residues to Glu
CC residues by deamination.
CC The enzyme is used in the prodn. of gelled foods, gelled cosmetics,
CC yogurt, gelatin, cheese etc. It is also used in the prodn. of
CC thermally stable materials such as microcapsules and carriers of
CC immobilized enzymes. The DNA sequence given allows the prodn. of
CC BTG efficiently and in large quantity.
XX
XX Sequence 406 AA;
SQ

```

Query Match 100.0%; Score 1811; DB 13; Length 406;
 Best Local Similarity 100.0%; Pred. No. 3e-153;
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSDRVTPPAELDRMPDPYPSYGRAETVNNYIRKQOQVYSHRDRGKQOQTEOREWL 60
 Db 76 dsddrvtppaelpdrmpdpypsygraetvnnnyirkwqvyshrdgrkqgmteedrewl 135
 QY 61 SYGCVGVTWVNSGYPTNRLAFASFEDEFRKNEKNGRPSGETRAEFEGRVAKESFDEE 120
 Db 136 sygcvgvtwvnsqgpyptnrlafasfdefrknelkngprsgetraefegrvakesfdee 195
 QY 121 KGFORAREVASVMNRALENHADESAYLDNKKELANGNDALRNEDARSPPYSALRNTPSF 180
 Db 196 kfgforarevasvmnralenahdesayldnkkelangndalrnedarspfysalrntpsf 255
 QY 181 KERNGNHDPSPRMKAVIYSKHFWSGQDRSSADKRYGDPDAFRPAGTGLVDMRSRDRNI 240
 Db 256 kerngnhdpssrmkaviyskhfwsqgdrssadkrkygdpdafrpapggtglvdmrdrni 315
 QY 241 PRSPSPGEGFVNFYDYGFGAQTADADKTWTHGNHYHAPNGSLGAMHYVESKFRNWE 300
 Db 316 prspspggefvnfydygfgaqtadadktvthgnhyhapngslgamhyveskfrnwse 375
 QY 301 GYSDFDRGAYVITFIPKSWNTAPDKVKQGW 331
 Db 376 gysdfdrgayvittfipkswntapdkvkqgwp 406

RESULT 9

AAB97831
 ID AAB97831 standard; Protein; 407 AA.

XX AAB97831;

DT 09-AUG-2001 (first entry)

DE S. mobaraense IFO 13819 transglutaminase protein SEQ ID NO:4.

XX Streptovorticillium cinnamonum IFO 12852; Streptomyces; actinomycete;

KW Streptovorticillium mobaraense IFO 13819; transglutaminase.

XX Streptovorticillium mobaraensis.

OS WO200129187-A1.

PN 26-APR-2001.

PD 13-OCT-2000; 2000WO-JP07135.

PF 18-OCT-1999; 99JP-0295649.

XX (AJIN) AJINOMOTO CO INC.

PA Taguchi S, Momose H;

PI WPI; 2001-300330/31.

XX N-PSDB; AAH20188.

XX Streptomyces sp. carrying an actinomycete-derived gene and promoter for

PT producing high yields of transglutaminase

PS Disclosure; Page 33-36; 41pp; Japanese.

XX The present invention describes a Streptomyces sp. containing a gene
 CC construct comprising actinomycete-derived transglutaminase gene and
 CC promoter. Also described are methods for producing pro-transglutaminase
 CC and active transglutaminase. The gene construct can be used in the
 CC production of large amounts of transglutaminase. The present sequence
 CC represents Streptovorticillium mobaraense IFO 13819 transglutaminase,
 CC which is given in the exemplification of the present invention.

XX Sequence 407 AA;

Query Match 100.0%; Score 1811; DB 22; Length 407;

Best Local Similarity 100.0%; Pred. No. 3e-153;

Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DSDRVTPPAELDRMPDPYPSYGRAETVNNYIRKQOQVYSHRDRGKQOQTEOREWL 60
 Db 76 dsddrvtppaelpdrmpdpypsygraetvnnnyirkwqvyshrdgrkqgmteedrewl 136
 QY 61 SYGCVGVTWVNSGYPTNRLAFASFEDEFRKNEKNGRPSGETRAEFEGRVAKESFDEE 120
 Db 137 sygcvgvtwvnsqgpyptnrlafasfdefrknelkngprsgetraefegrvakesfdee 196
 QY 121 KGFORAREVASVMNRALENHADESAYLDNKKELANGNDALRNEDARSPPYSALRNTPSF 180
 Db 197 kfgforarevasvmnralenahdesayldnkkelangndalrnedarspfysalrntpsf 256
 QY 181 KERNGNHDPSPRMKAVIYSKHFWSGQDRSSADKRYGDPDAFRPAGTGLVDMRSRDRNI 240
 Db 257 kerngnhdpssrmkaviyskhfwsqgdrssadkrkygdpdafrpapggtglvdmrdrni 316
 QY 241 PRSPSPGEGFVNFYDYGFGAQTADADKTWTHGNHYHAPNGSLGAMHYVESKFRNWE 300
 Db 317 prspspggefvnfydygfgaqtadadktvthgnhyhapngslgamhyveskfrnwse 376
 QY 301 GYSDFDRGAYVITFIPKSWNTAPDKVKQGW 331
 Db 377 gysdfdrgayvittfipkswntapdkvkqgwp 407

RESULT 10

AAB81166

ID AAB81166 standard; Protein; 407 AA.

XX AAB81166;

DT 13-JUL-2001 (first entry)

DE Prepro-transglutaminase amino acid sequence.

KW Coryneform bacteria; transglutaminase; food processing.

XX Streptovorticillium mobaraense.

OS WO200123591-A1.

PN 05-APR-2001.

PD 29-SEP-2000; 2000WO-JP06780.

PF 30-SEP-1999; 99JP-0280098.

XX 28-JUN-2000; 2000JP-0194043.

PA (AJIN) AJINOMOTO CO INC.

PI Kikuchi Y, Date M, Umezawa Y, Yokoyama K, Matsui H;

XX WPI; 2001-266172/27.

XX N-PSDB; AAF86283.

XX Efficient secretory production of foreign proteins e.g.
 PT transglutaminase employing transformant coryneform bacterium, simply on
 PT industrial scale with direct recovery for use in food processing and
 PT pharmaceutical industry

PS Example 1; Page 90-93; 151pp; Japanese.

XX This invention relates to a process for the production of a foreign
 CC secretory protein through the construction of a recombinant coryneform
 CC bacterium. The coryneform bacterium is transformed with an expression
 CC construct in which DNA encoding a target foreign protein pro-structure is
 CC ligated to the downstream region of DNA encoding the signal peptide
 CC domain of a coryneform bacterial protein. Following transformation with
 CC the vector, the bacterium is cultured, and the pro-peptide cleaved from
 CC the expressed protein. Transglutaminases produced using this process are
 CC useful in the food processing and pharmaceutical industries. The present

CC sequence prepro-transglutaminase. The protein is used in an example
 CC illustrating the method of the invention.

SQ Sequence 407 AA;

Query Match 100.0%; Score 1811; DB 22; Length 407;
 Best Local Similarity 100.0%; Pred. No. 3e-153;
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSDRVTTPAEPLDRMPDPRPSYGRAETVYNNYIRKWOQVYSHRDGRKQKQMTQEQRWL 60
 DB 77 dsdrrvtppaepldrmpdprpsygraetvnnnyirkwqvshrdgrkqgnteeqrwl 136
 QY 61 SYGCVGTWVNSGYPTNRLAFASDFEDFRKNEKNGRPSGSETRAEFGRVAKESFDEE 120
 DB 137 sygcvgvtwnsgyptnrlafasdfedfrknelkngprsgsetraefgrvakesfdee 196
 QY 121 KGFQRAREVASVMNRALENAHDESAYLDNLKKELANGNDALRNEDARSFPYSALRNTPSF 180
 DB 197 kgfqrarevasvmnralenahdesayldnlkkelangndalrnedarspfysalrntpsf 256
 QY 181 KERNGGNHDPSPRMKAVIYISKHFWSGQDRSSADKKRYGDPDAFRPAPGTGLVDMSRDRNI 240
 DB 257 kernggnhdpssrmkaviyskhfwsqdrssadkrygdpdafpapggtglvdmrdrni 316
 QY 241 PRSPTSPGSEGVNFYDYGWFGAQTADADKTWTHGNHYHAPNGSLGAMHYVESKFRNWSE 300
 DB 317 prsptspgegvnfdygwfgaqteadadktwthgnhyhapngslgamhyveskfrnwse 376
 QY 301 GYSDFDRGAYVITFIPKSWNTAPDKVKQGW 331
 DB 377 gysdfdrgayvitfipkswntapdkvkqgw 407

RESULT 11

AAR49049
 ID AAR49049 standard; Protein; 408 AA.

AC AAR49049;

DT 20-SEP-1994 (first entry)

XX BTG-contg. sequence.

XX Bacterial transglutaminase; BTG; expression;
 KW active; inactive; inclusion body.

XX JP06030771-A.

XX 08-FEB-1994.

XX 14-JUL-1992; 92JP-0187038.

XX 14-JUL-1992; 92JP-0187038.

XX (AJIN) AJINOMOTO KK.

XX WPI; 1994-079294/10.

XX N-PSDB; AA055984.

XX Prepn. of bacterial trans-glutaminase in large amts. - by

PT expression of fusion protein in E. coli bacterial

PT trans-glutaminase

XX Disclosure; Page 8-10; 13pp; Japanese.

XX A fused protein contains amino acids 16-346 of BTG (AAQ55983)
 CC and a hydrophilic peptide at the amino terminal.
 CC Expression of DNA encoding this protein in E. coli allow large
 CC scale prodn. of BTG. An active BTG can be prepd. from the
 CC inactive fused protein inclusion body.

XX

SQ Sequence 408 AA;

Query Match 100.0%; Score 1811; DB 15; Length 408;
 Best Local Similarity 100.0%; Pred. No. 3e-153;
 Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSDRVTTPAEPLDRMPDPRPSYGRAETVYNNYIRKWOQVYSHRDGRKQKQMTQEQRWL 60
 DB 78 dsdrrvtppaepldrmpdprpsygraetvnnnyirkwqvshrdgrkqgnteeqrwl 137
 QY 61 SYGCVGTWVNSGYPTNRLAFASDFEDFRKNEKNGRPSGSETRAEFGRVAKESFDEE 120
 DB 138 sygcvgvtwnsgyptnrlafasdfedfrknelkngprsgsetraefgrvakesfdee 197
 QY 121 KGFQRAREVASVMNRALENAHDESAYLDNLKKELANGNDALRNEDARSFPYSALRNTPSF 180
 DB 198 kgfqrarevasvmnralenahdesayldnlkkelangndalrnedarspfysalrntpsf 257
 QY 181 KERNGGNHDPSPRMKAVIYISKHFWSGQDRSSADKKRYGDPDAFRPAPGTGLVDMSRDRNI 240
 DB 258 kernggnhdpssrmkaviyskhfwsqdrssadkrygdpdafpapggtglvdmrdrni 317
 QY 241 PRSPTSPGSEGVNFYDYGWFGAQTADADKTWTHGNHYHAPNGSLGAMHYVESKFRNWSE 300
 DB 318 prsptspgegvnfdygwfgaqteadadktwthgnhyhapngslgamhyveskfrnwse 377
 QY 301 GYSDFDRGAYVITFIPKSWNTAPDKVKQGW 331
 DB 378 gysdfdrgayvitfipkswntapdkvkqgw 408

RESULT 12

AAR22653
 ID AAR22653 standard; Protein; 331 AA.

AC AAR22653;

DT 09-OCT-1992 (first entry)

XX Transglutaminase (expressed in E. coli).

XX BTG; acyl rearrangement; deamination; yeast; actinomycetes.

XX Synthetic.

XX EP481504-A.

XX 22-APR-1992.

XX 18-OCT-1991; 91EP-0117813.

XX 19-OCT-1990; 90JP-0282566.

XX (AJIN) AJINOMOTO CO INC.

XX (AMANO) AMANO PHARM KK.

XX Ando K, Arafuka S, Koikeda S, Matsui H, Takagi H, Washizu K;

XX WPI; 1992-133808/17.

XX N-PSDB; AAQ24207.

XX DNA fragment encoding trans:glutaminase - is inserted into

PT vector, e.g. pNJI053-BTG, for protein expression

XX Disclosure; Page 3; 55pp; English.

XX The protein sequence given has transglutaminase (BTG) activity. The
 CC DNA encoding this protein has a base sequence which can be used
 CC suitably in an expression system using E. coli or yeast as a host.
 CC The base sequence can be compared to those given in AAQ24197 and
 CC AAQ24200 which are derived from actinomycetes sp.
 CC BTG catalyses an acyl rearrangement reaction of a gamma-carboxamide

CC gp. of glutamine. It introduces intra- or intermolecular formation
 CC of epsilon-(gamma-Gln)-Lys cross-linking when an epsilon-amino gp.
 CC of a Lys residue acts as an acyl receptor. When water acts as an
 CC acyl acceptor the enzyme accelerates the conversion of Gln residues
 CC to Gln residues by deamination.
 CC The enzyme is used in the prodn. of gelled foods, gelled cosmetics,
 CC yogurt, gelatin, cheese etc. It is also used in the prodn. of
 CC thermally stable materials such as microcapsules and carriers of
 CC immobilized enzymes. The DNA sequence given allows the prodn. of
 CC BFG efficiently and in large quantity.
 XX Sequence 331 AA;
 SQ

Query Match 99.5%; Score 1802; DB 13; Length 331;
 Best Local Similarity 99.7%; Pred. No. 1.4e-152;
 Matches 330; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 DSDDRVTPAEPLDRMPDPYRPSYGRAETVNNYIRKQOQVYSHRDKRQOQTEQREWL 60
 Dd 1 dddrvtpaepldrmpdpypysygraetvnnnyirkwqvysyhrdgrkqgmteeqrewl 60

Qy 61 SYGCVGVTVWNSGQYPTNRLAFASDEDRFKNELKNGRPSGTRAEFEGRVAKESFDEE 120
 Dd 1 dddrvtpaepldrmpdpypysygraetvnnnyirkwqvysyhrdgrkqgmteeqrewl 60

Qy 61 SYGCVGVTVWNSGQYPTNRLAFASDEDRFKNELKNGRPSGTRAEFEGRVAKESFDEE 120
 Dd 1 dddrvtpaepldrmpdpypysygraetvnnnyirkwqvysyhrdgrkqgmteeqrewl 60

Qy 121 KGFQREAVASVMNRALENAHDESAYLDNLKELANGNDALRNEDARSPFYSALENTPSF 180
 Dd 121 kfgfqravvasvmnralenahdesayldnlkkelangndalrnedarspfysalrntpsf 180

Qy 181 KERNGNHPDSRMKAVIYSKHFWSGQDRSSADKRYGDDPDAFRPAGTGLVDMSRDRI 240
 Dd 181 kerngnchpdsrmkaviyskhfwsgdrrssadkrygdpdafrpaptglvdmrsvrni 240

Qy 241 PRSPTSGGCFVNFYDYGWFGAQTADADKTYVTHGNHHPNGSLGAMHYVESKFRNWE 300
 Dd 241 prsptspegcfvnyfydgwfgaqteadadktvthghnhpangslgamhvyeskrnwse 300

Qy 301 GYSDFORGAVITFIPKSWNTAPDKVKQGP 331
 Dd 301 gysdforgavvitfipkswntapdkvkqgp 331

RESULT 13
 AAB47007
 ID AAB47007 standard; Protein; 400 AA.
 AC AAB47007;
 XX
 DT 22-MAR-2001 (first entry)
 XX
 DE Transglutaminase.
 XX
 KW Transglutaminase; acyl transfer; glutamine; ATCC 27446; food;
 KW cosmetic; pharmaceutical; gel; artificial skin.
 XX
 OS Streptococcus thermophilus
 XX
 PN W0200070026-A1.
 XX
 PD 23-NOV-2000.
 XX
 PF 10-MAY-2000; 2000WO-US12601.
 XX
 PR 14-MAY-1999; 99US-0134158.
 XX
 XX (WISC) WISCONSIN ALUMNI RES FOUND.
 XX Damodaran S;
 XX
 PI WPI: 2001-070774/08.
 DR N-PSDB; AAC85251.
 XX

PT New transglutaminase enzyme from Streptococcus thermophilus
 PT e.g. use in gelled products and the production of artificial skin
 XX Claim 8; Page 44-45; 49pp; English.
 CC This sequence represents a transglutaminase which catalyzes the acyl
 CC transfer of the gamma-carboxamide group of a glutamine residue in
 CC a peptide or protein chain independently of calcium ions and which
 CC has an activity at pH 9.0 that is 40 % or greater than its activity
 CC at pH 7.0. The transglutaminase was isolated from Streptococcus
 CC thermophilus strain ATCC (American Type Culture Collection) No. 27446.
 CC The transglutaminase catalyzes an acyl transfer reaction of a
 CC gamma-carboxamide group of a glutamine residue and a primary amine of
 CC a peptide. When the epsilon-amino group of a lysine residue functions as
 CC the acyl acceptor, intramolecular and intermolecular cross-linking
 CC occurs. When water functions as the acyl acceptor, transglutaminase
 CC converts glutamine residues in glutamic acid residues by deamidation.
 CC The cross-linking reaction is useful in the food, cosmetic and
 CC pharmaceutical industries. Transglutaminase can gel protein, making it
 CC useful in production of gelled food, gelled cosmetics, gelatins,
 CC yogurt, cheese and other products. The enzyme can also be used to make
 CC thermally stable materials such as microcapsules or carriers of
 CC immobilized enzymes. The cross-linking reaction is also potentially
 CC useful in production of artificial skin. The transglutaminase coding
 CC sequence may be used as probes for detecting the presence and/or
 CC expression of streptococcus transglutaminase genes, or for
 CC identifying related genes from other microbial species. They may
 CC also be used to produce large quantities of the enzyme.
 CC The new transglutaminase has a higher activity, both at pH
 CC 7.0 and pH 9.0, when compared to commercially available
 CC transglutaminases. It has a different susceptibility to several
 CC commonly used inhibitors, when compared to previously isolated
 CC transglutaminases. These properties broaden the range of application
 CC in which the new enzyme can be used.
 XX Sequence 400 AA;
 SQ

Query Match 82.4%; Score 1493; DB 22; Length 400;
 Best Local Similarity 82.3%; Pred. No. 7.3e-125;
 Matches 269; Conservative 25; Mismatches 33; Indels 0; Gaps 0;

Qy 5 RVTPAEPLDRMPDPYRPSYGRAETVNNYIRKQOQVYSHRDKRQOQTEQREWLISYGC 64
 Dd 74 retppaepldrmpdayrarggrattvnnnyirkwqvysyhrdgrkqgmteeqreklisygc 133

Qy 65 VGVTVWNSGQYPTNRLAFASDEDRFKNELKNGRPSGTRAEFEGRVAKESFDEEGFQ 124
 Dd 134 vgvtnvnsqpyptnrlafsfddedkykndlkntrplagetratrefgriakafdegkfk 193

Qy 125 RAREVASVMNRALENAHDESAYLDNLKELANGNDALRNEDARSPFYSALENTPSFERN 184
 Dd 194 rardvaslmmkalenahdegayldklxtelnnndalhedrsnfysalrntpsferrd 253

Qy 185 GGNHDPSSRMKAVIYSKHFWSGQDRSSADKRYGDDPDAFRPAGTGLVDMSRDRI 244
 Dd 254 ggnydpsrmkaviyskhfwsgdrgssdkrygdpdafrpdrgrtglvdmkskdnirpsp 313

Qy 245 TSPGEGFVNFYDYGWFGAQTADADKTYVTHGNHHPNGSLGAMHYVESKFRNWE 304
 Dd 314 anpgegwnfydgwfgaqteadadktvthghnhpangdlgpmhvyeskrfnwsagyd 373

Qy 305 FDRGAVITFIPKSWNTAPDKVKQGP 331
 Dd 374 fdrgtymiallpkswntapdkvkqgp 400

RESULT 14
 AAB81164
 ID AAB81164 standard; protein; 330 AA.
 XX
 AC AAB81164;
 XX

DT 13-JUL-2001 (first entry)
 XX Transglutaminase related protein SEQ ID 43.
 DE Coryneform bacteria; transglutaminase; food processing.
 XX Streptovorticillum cinnamomeum.
 XX WO200123591-A1.
 PN 05-APR-2001.
 XX 29-SEP-2000; 2000WO-JP06780.
 XX 30-SEP-1999; 99JP-0280098.
 PR 28-JUN-2000; 2000JP-0194043.
 XX (AJIN) AJINOMOTO CO INC.
 PA Kikuchi Y, Date M, Umezawa Y, Yokoyama K, Matsui H;
 PI WPI; 2001-256172/27.
 XX Efficient secretory production of foreign proteins e.g.
 PT transglutaminase employing transformant coryneform bacterium, simply on
 PT industrial scale with direct recovery for use in food processing and
 PT pharmaceutical industry -
 XX Claim 24; Page 131-133; 151pp; Japanese.
 XX This invention relates to a process for the production of a foreign
 CC secretory protein through the construction of a recombinant coryneform
 CC bacterium. The coryneform bacterium is transformed with an expression
 CC construct in which DNA encoding a target foreign protein pro-structure is
 CC ligated to the downstream region of DNA encoding the signal peptide
 CC domain of a coryneform bacterial protein. Following transformation with
 CC the vector, the bacterium is cultured and the propeptide cleaved from
 CC the expressed protein. Transglutaminases produced using this process are
 CC useful in the food processing and pharmaceutical industries. The present
 CC sequence represents a transglutaminase related protein, which can be
 CC used in the method of the invention.
 XX Sequence 330 AA;
 SQ
 Query Match 82.1%; Score 1487; DB 22; Length 330;
 Best Local Similarity 81.5%; Pred. No. 1.9e-124;
 Matches 269; Conservative 26; Mismatches 35; Indels 0; Gaps 0;
 QY 2 SDDRVTTPAEPDLRMPDPYRPSYGRAETVNNYIRKWOQVYSHRDGRKQOQTEEQREWLS 61
 DB 1 sddretppaepdlrmpdayraygrattvnnnyirkwqvyshrdgkkgqmtteeqrekl 60
 QY 62 YGCVGVTVWNSGOYPTNRLAFASFDEDFRKNELKNGRPSRGETRAEFGRVAKESFDEEK 121
 DB 61 ygcvgvtwvnsqgpyptnrlafasfdenkyndkntsprdpdtraefegriakgsfdegk 120
 QY 122 GFORAREVASVMNRALENAHDEGSAYLDNLKELANGNDALRNEDARSPPFYSALRNTPSFK 181
 DB 121 gfkardvasvmnkalenahdegtyinnlktelntnndallredsrnsfysalrntpsfk 180
 QY 182 ERNGGNHDPGRMKAVIYSKHFWSGQDRSSADRKRYGDPDAFPAPGCTGLVDMRSRNP 241
 DB 181 erdggnydpkkmkavlyskhfwsgdqrgssdkrykypdpeafpdpdgtglvdmksdrsp 240
 QY 242 RSPTSPGCGFVNFYDYGWFGAQTEADADKTVTHGNHYHAPNGSLGAMHVEYKSKFRNWSG 301
 DB 241 rspakpgegvwnfydgwfgaqteadadktwtghdghyhapnsdlgpmhvnheskfrkwsag 300
 QY 302 YSDFDRGAYVITFIPKSWNTAPDKVKQGW 331
 DB 301 yadfdrgayvitfipkswntapakveggwp 330

RESULT 15
 AAB97830
 ID AAB97830 standard; Protein; 416 AA.
 XX
 AC AAB97830;
 XX 09-AUG-2001 (first entry)
 DT
 DE S. cinnamomeum IFO 12852 transglutaminase protein SEQ ID NO:2.
 XX Streptovorticillum cinnamomeum IFO 12852; Streptomyces; actinomycete;
 KW Streptovorticillum mobarsense IFO 13819; transglutaminase.
 XX Streptovorticillum cinnamomeus.
 OS
 XX WO200129187-A1.
 PN 26-APR-2001.
 PD 13-OCT-2000; 2000WO-JP07135.
 PF 18-OCT-1999; 99JP-0295649.
 PR (AJIN) AJINOMOTO CO INC.
 PA Taguchi S, Momose H;
 PI WPI; 2001-300330/31.
 DR N-PSDB; AAB20187.
 XX Streptomyces sp. carrying an actinomycete-derived gene and promoter for
 PT producing high yields of transglutaminase -
 CC Claim 6; Page 26-28; 41pp; Japanese.
 XX The present invention describes a Streptomyces sp. containing a gene
 CC construct comprising actinomycete-derived transglutaminase gene and
 CC promoter. Also described are methods for producing pro-transglutaminase
 CC and active transglutaminase. The gene construct can be used in the
 CC production of large amounts of transglutaminase. The present sequence
 CC represents Streptovorticillum cinnamomeum IFO 12852 transglutaminase,
 CC which is used in the present invention.
 XX Sequence 416 AA;
 SQ
 Query Match 82.1%; Score 1487; DB 22; Length 416;
 Best Local Similarity 81.5%; Pred. No. 2.6e-124;
 Matches 269; Conservative 26; Mismatches 35; Indels 0; Gaps 0;
 QY 2 SDDRVTTPAEPDLRMPDPYRPSYGRAETVNNYIRKWOQVYSHRDGRKQOQTEEQREWLS 61
 DB 87 sddretppaepdlrmpdayraygrattvnnnyirkwqvyshrdgkkgqmtteeqrekl 146
 QY 62 YGCVGVTVWNSGOYPTNRLAFASFDEDFRKNELKNGRPSRGETRAEFGRVAKESFDEEK 121
 DB 147 ygcvgvtwvnsqgpyptnrlafasfdenkyndkntsprdpdtraefegriakgsfdegk 206
 QY 122 GFORAREVASVMNRALENAHDEGSAYLDNLKELANGNDALRNEDARSPPFYSALRNTPSFK 181
 DB 207 gfkardvasvmnkalenahdegtyinnlktelntnndallredsrnsfysalrntpsfk 266
 QY 182 ERNGGNHDPGRMKAVIYSKHFWSGQDRSSADRKRYGDPDAFPAPGCTGLVDMRSRNP 241
 DB 267 erdggnydpkkmkavlyskhfwsgdqrgssdkrykypdpeafpdpdgtglvdmksdrsp 326
 QY 242 RSPTSPGCGFVNFYDYGWFGAQTEADADKTVTHGNHYHAPNGSLGAMHVEYKSKFRNWSG 301
 DB 327 rspakpgegvwnfydgwfgaqteadadktwtghdghyhapnsdlgpmhvnheskfrkwsag 386
 QY 302 YSDFDRGAYVITFIPKSWNTAPDKVKQGW 331
 DB 302 yadfdrgayvitfipkswntapakveggwp 330

Db 387 yadfrgavvtfipkswntapakveggwp 416

Search completed: September 27, 2002, 12:39:35
Job time: 88 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2002, 12:38:07 ; Search time 13.02 Seconds
(without alignments)
620.958 Million cell updates/sec

Title: US-09-884-948-1
Perfect score: 1811
Sequence: 1 DSDRVPFAPELDRMPDY.....ITFIPKSWNTAPDKVKQGP 331
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents_AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1811	100.0	331	1	US-08-136-993-1
2	1811	100.0	331	3	US-09-109-063-1
3	1811	100.0	331	3	US-08-793-426A-3
4	1811	100.0	331	4	US-09-294-565-3
5	1811	100.0	406	1	US-08-136-993-13
6	1465	80.9	331	3	US-08-793-426A-2
7	1465	80.9	331	4	US-09-294-565-2
8	591	32.6	126	3	US-08-793-426A-7
9	591	32.6	126	4	US-09-294-565-7
10	208	11.5	40	3	US-08-793-426A-8
11	208	11.5	40	4	US-09-294-565-8
12	189	10.4	52	3	US-08-793-426A-6
13	189	10.4	52	4	US-09-294-565-6
14	99	5.5	315	3	US-08-558-135-5
15	91	5.0	544	1	US-08-264-002-7
16	90.5	5.0	367	3	US-09-041-889-30
17	87.5	4.8	1128	1	US-08-111-939-2
18	86.5	4.8	913	1	US-08-487-890A-5
19	86.5	4.8	913	2	US-08-478-435-5
20	86.5	4.8	913	3	US-08-337-483-5
21	86.5	4.8	913	2	US-08-478-373-5
22	86.5	4.8	913	3	US-08-474-671-5
23	86.5	4.8	913	3	US-08-483-577A-5
24	86.5	4.8	913	4	US-08-897-438-5
25	86.5	4.8	913	4	US-08-637-654-5
26	84.5	4.7	625	4	US-09-360-197-15
27	84.5	4.7	674	3	US-08-893-852A-1

28 84 4.6 453 4 US-08-961-083-38
29 84 4.6 535 3 US-08-725-459B-18
30 84 4.6 630 3 US-08-725-459B-17
31 84 4.6 680 3 US-08-725-459B-3
32 84 4.6 730 3 US-08-725-459B-2
33 84 4.6 733 3 US-08-725-459B-23
34 84 4.6 733 3 US-08-725-459B-28
35 84 4.6 733 3 US-08-725-459B-29
36 84 4.6 733 3 US-08-725-459B-30
37 84 4.6 748 3 US-08-725-459B-24
38 84 4.6 748 3 US-08-725-459B-27
39 84 4.6 753 3 US-08-725-459B-25
40 84 4.6 753 3 US-08-725-459B-26
41 84 4.6 769 3 US-08-725-459B-35
42 84 4.6 769 3 US-08-725-459B-36
43 84 4.6 769 3 US-08-725-459B-37
44 84 4.6 769 3 US-08-725-459B-38
45 84 4.6 769 3 US-08-725-459B-39

ALIGNMENTS

RESULT 1
US-08-136-993-1
; Sequence 1, Application US/08136993
; Patent No. 5420025
; GENERAL INFORMATION:
; APPLICANT: Takagi, Hiroshi
; APPLICANT: Arafuka, Shino
; APPLICANT: Matsui, Hiroshi
; APPLICANT: Washizu, Kinya
; APPLICANT: Ando, Keiichi
; APPLICANT: Koikeda, Satoshi
; TITLE OF INVENTION: Recombinant transglutaminase
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue
; CITY: N.W.
; STATE: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/136.993
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/777.447
; FILING DATE:
; APPLICATION NUMBER: JP 2-282566
; FILING DATE: 19-OCT-1990
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-293-7060
; TELEFAX: 202-293-7860
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-136-993-1

Query Match 100.0%; Score 1811; DB 1; Length 331;
Best Local Similarity 100.0%; Pred. No. 3.8e-171;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSDDRVTPPAELDRMPDPYRPSYGRAETVYNNYIRKWOQVYSHRDGRKQOQTEQREWL 60
DB 1 DSDDRVTPPAELDRMPDPYRPSYGRAETVYNNYIRKWOQVYSHRDGRKQOQTEQREWL 60
QY 61 SYGCVGVTVWNSGOYPTNRLAFASDEDFRKNELKNGRPSGTRAEFEGRVAKESFDEE 120
DB 61 SYGCVGVTVWNSGOYPTNRLAFASDEDFRKNELKNGRPSGTRAEFEGRVAKESFDEE 120
QY 121 KGFQREVASVNNRALENNAHDESAVLDNLKELANGNDALRNEDARSFFYSALRNTPSF 180
DB 121 KGFQREVASVNNRALENNAHDESAVLDNLKELANGNDALRNEDARSFFYSALRNTPSF 180
QY 181 KERNGNHDPDSRMKAVIYSKHFWSGQDRSSADKRYGDPDAFRPAPGTGLVDMSDRNI 240
DB 181 KERNGNHDPDSRMKAVIYSKHFWSGQDRSSADKRYGDPDAFRPAPGTGLVDMSDRNI 240
QY 241 PRSPTSPGEGFVNFYDYGWGAQTEADADKTWTHGNHYHAPNGSLGAMHVYESKFRNWSE 300
DB 241 PRSPTSPGEGFVNFYDYGWGAQTEADADKTWTHGNHYHAPNGSLGAMHVYESKFRNWSE 300
QY 301 GYSDFDRGAYVITFIPKSWNTAPDKVKQGW 331
DB 301 GYSDFDRGAYVITFIPKSWNTAPDKVKQGW 331

RESULT 2

US-09-109-063-1
; Sequence 1, Application US/09109063
; Patent No. 6013498
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, KEIICHI
; APPLICANT: NAKAMURA, NAMI
; APPLICANT: MIWA, TETSUYA
; APPLICANT: SEGURO, KATSUYA
; TITLE OF INVENTION: PROCESS FOR PRODUCING MICROBIAL TRANSGLUTAMINASE
; FILE REFERENCE: 0010-0937-0
; CURRENT APPLICATION NUMBER: US/09/109,063
; CURRENT FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: JP 180010/1997
; EARLIER FILING DATE: 1997-07-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial
; OTHER INFORMATION: Sequence: TRANSGLUTAMINASE
US-09-109-063-1

Query Match 100.0%; Score 1811; DB 3; Length 331;
Best Local Similarity 100.0%; Pred. No. 3, 8e-171;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSDDRVTPPAELDRMPDPYRPSYGRAETVYNNYIRKWOQVYSHRDGRKQOQTEQREWL 60
DB 1 DSDDRVTPPAELDRMPDPYRPSYGRAETVYNNYIRKWOQVYSHRDGRKQOQTEQREWL 60
QY 61 SYGCVGVTVWNSGOYPTNRLAFASDEDFRKNELKNGRPSGTRAEFEGRVAKESFDEE 120
DB 61 SYGCVGVTVWNSGOYPTNRLAFASDEDFRKNELKNGRPSGTRAEFEGRVAKESFDEE 120
QY 121 KGFQREVASVNNRALENNAHDESAVLDNLKELANGNDALRNEDARSFFYSALRNTPSF 180
DB 121 KGFQREVASVNNRALENNAHDESAVLDNLKELANGNDALRNEDARSFFYSALRNTPSF 180
QY 181 KERNGNHDPDSRMKAVIYSKHFWSGQDRSSADKRYGDPDAFRPAPGTGLVDMSDRNI 240
DB 181 KERNGNHDPDSRMKAVIYSKHFWSGQDRSSADKRYGDPDAFRPAPGTGLVDMSDRNI 240

QY 241 PRSPTSPGEGFVNFYDYGWGAQTEADADKTWTHGNHYHAPNGSLGAMHVYESKFRNWSE 300
DB 241 PRSPTSPGEGFVNFYDYGWGAQTEADADKTWTHGNHYHAPNGSLGAMHVYESKFRNWSE 300
QY 301 GYSDFDRGAYVITFIPKSWNTAPDKVKQGW 331
DB 301 GYSDFDRGAYVITFIPKSWNTAPDKVKQGW 331

RESULT 3

US-08-793-426A-3
; Sequence 3, Application US/08793426A
; Patent No. 6100053
; GENERAL INFORMATION:
; APPLICANT: Bech, Lisbeth
; APPLICANT: No. 6100053revang, Iben
; APPLICANT: Halkier, Torben
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Schafer, Thomas
; APPLICANT: Andersen, Jens
; TITLE OF INVENTION: Microbial Transglutaminases, Their
; TITLE OF INVENTION: Production And Use
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6100053o No. 6100053disk of No. 6100053th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,426A
; FILING DATE: 25-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol E.
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 4211.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-793-426A-3

Query Match 100.0%; Score 1811; DB 3; Length 331;
Best Local Similarity 100.0%; Pred. No. 3, 8e-171;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSDDRVTPPAELDRMPDPYRPSYGRAETVYNNYIRKWOQVYSHRDGRKQOQTEQREWL 60
DB 1 DSDDRVTPPAELDRMPDPYRPSYGRAETVYNNYIRKWOQVYSHRDGRKQOQTEQREWL 60
QY 61 SYGCVGVTVWNSGOYPTNRLAFASDEDFRKNELKNGRPSGTRAEFEGRVAKESFDEE 120
DB 61 SYGCVGVTVWNSGOYPTNRLAFASDEDFRKNELKNGRPSGTRAEFEGRVAKESFDEE 120
QY 121 KGFQREVASVNNRALENNAHDESAVLDNLKELANGNDALRNEDARSFFYSALRNTPSF 180
DB 121 KGFQREVASVNNRALENNAHDESAVLDNLKELANGNDALRNEDARSFFYSALRNTPSF 180

QY 181 KERNGNHDPSRMKAVIYSKHFWSGQDRSSSADKRYGDPDAFRPAPGTGLVDMSSDRNI 240
|||||
Db 181 KERNGNHDPSRMKAVIYSKHFWSGQDRSSSADKRYGDPDAFRPAPGTGLVDMSSDRNI 240
|||||
QY 241 PRSPTSPGEGFVNFYDYGWFGAQTEADADKTYWTHGNHYHAPNGSLGAMHVYESKFRNWSE 300
|||||
Db 241 PRSPTSPGEGFVNFYDYGWFGAQTEADADKTYWTHGNHYHAPNGSLGAMHVYESKFRNWSE 300
|||||
QY 301 GYSDFDRGAYVITFIPKSWNTAPDKVKGWP 331
|||||
Db 301 GYSDFDRGAYVITFIPKSWNTAPDKVKGWP 331
|||||
RESULT 4
US-09-294-565-3
; Sequence 3, Application US/09294565
; Patent No. 6190879
; GENERAL INFORMATION:
; APPLICANT: Bech, Lisbeth
; APPLICANT: No. 6190879revang, Iben
; APPLICANT: Halkier, Torben
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Schafer, Thomas
; APPLICANT: Andersen, Jens
; TITLE OF INVENTION: Microbial Transglutaminases, Their
; TITLE OF INVENTION: Production And Use
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 61908790 No. 6190879disk of No. 6190879th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/294,565
; FILING DATE: 19-APR-1999
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4211.224-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-294-565-3

Query Match 100.0%; Score 1811; DB 4; Length 331;
Best Local Similarity 100.0%; Pred. No. 3,8e-171;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSDDRVTTPAEPDLRMPDPYRPSYGRAETVNNYIRKQOQVYSHRDGRKQOQTEOREWL 60
|||||
Db 1 DSDDRVTTPAEPDLRMPDPYRPSYGRAETVNNYIRKQOQVYSHRDGRKQOQTEOREWL 60
|||||
QY 61 SYGCGVGTWVNSGQVPTNRLAFASDEDRFKNELKNGRPSRGETRAEFEGRVAKESFDEE 120
|||||
Db 61 SYGCGVGTWVNSGQVPTNRLAFASDEDRFKNELKNGRPSRGETRAEFEGRVAKESFDEE 120
|||||

QY 121 KGFQREAVSVNRALENHDSAYLDNLKKELANGNDALRNEDARSPPYSALRNTPSF 180
|||||
Db 121 KGFQREAVSVNRALENHDSAYLDNLKKELANGNDALRNEDARSPPYSALRNTPSF 180
|||||
QY 181 KERNGNHDPSRMKAVIYSKHFWSGQDRSSSADKRYGDPDAFRPAPGTGLVDMSSDRNI 240
|||||
Db 181 KERNGNHDPSRMKAVIYSKHFWSGQDRSSSADKRYGDPDAFRPAPGTGLVDMSSDRNI 240
|||||
QY 241 PRSPTSPGEGFVNFYDYGWFGAQTEADADKTYWTHGNHYHAPNGSLGAMHVYESKFRNWSE 300
|||||
Db 241 PRSPTSPGEGFVNFYDYGWFGAQTEADADKTYWTHGNHYHAPNGSLGAMHVYESKFRNWSE 300
|||||
QY 301 GYSDFDRGAYVITFIPKSWNTAPDKVKGWP 331
|||||
Db 301 GYSDFDRGAYVITFIPKSWNTAPDKVKGWP 331
|||||
RESULT 5
US-08-136-993-13
; Sequence 13, Application US/08136993
; Patent No. 5420025
; GENERAL INFORMATION:
; APPLICANT: Takagi, Hiroshi
; APPLICANT: Arafuka, Shinobu
; APPLICANT: Matsui, Hiroshi
; APPLICANT: Washizu, Kinya
; APPLICANT: Ando, Keiichi
; APPLICANT: Koikeda, Satoshi
; TITLE OF INVENTION: Recombinant transglutaminase
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
; STREET: 2100 Pennsylvania Avenue
; CITY: N.W.
; STATE: Washington, D.C.
; COUNTRY: U.S.A.
; ZIP: 20037-3202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/136,993
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/777,447
; FILING DATE:
; APPLICATION NUMBER: JP 2-282566
; FILING DATE: 19-OCT-1990
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-293-7060
; TELEFAX: 202-293-7360
; TELEX: 6491103
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 406 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-136-993-13

Query Match 100.0%; Score 1811; DB 1; Length 406;
Best Local Similarity 100.0%; Pred. No. 5.1e-171;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DSDDRVTTPAEPDLRMPDPYRPSYGRAETVNNYIRKQOQVYSHRDGRKQOQTEOREWL 60
|||||
Db 76 DSDDRVTTPAEPDLRMPDPYRPSYGRAETVNNYIRKQOQVYSHRDGRKQOQTEOREWL 135
|||||
QY 61 SYGCGVGTWVNSGQVPTNRLAFASDEDRFKNELKNGRPSRGETRAEFEGRVAKESFDEE 120
|||||

```

Db   136 SYGCVGVTVNWSQYPTNRLAFASDEDRFKNELKNRPSGETRAEFGRVAKESFDEE 195
      |||
Qy   121 KGFQREAVSYVMNRALENAHDESAYLDNLKKELANGNDALRNEDARSFPYSALRNTPSF 180
      |||
Db   196 KGFQREAVSYVMNRALENAHDESAYLDNLKKELANGNDALRNEDARSFPYSALRNTPSF 255
      |||
Qy   181 KERNGNHDPSPRMKAVIYSKHFWSCQDRSSADRRKYGDPAFPAPCTGLVDMSRDRI 240
      |||
Db   256 KERNGNHDPSPRMKAVIYSKHFWSCQDRSSADRRKYGDPAFPAPCTGLVDMSRDRI 315
      |||
Qy   241 PRSFTSPGEGFYNFYDYGFGAGTEADADKTVMTHGNHYHAPNGLSGAMHYVESKERNWSE 300
      |||
Db   316 PRSFTSPGEGFYNFYDYGFGAGTEADADKTVMTHGNHYHAPNGLSGAMHYVESKERNWSE 375
      |||
Qy   301 GYSDFDRGAVITFIPKSWNTAPDKVKOGWP 331
      |||
Db   376 GYSDFDRGAVITFIPKSWNTAPDKVKOGWP 406
      |||

RESULT        6
US-08-793-426A-2
; Sequence 2, Application US/08793426A
; Patent No. 6100053
; GENERAL INFORMATION:
; APPLICANT: Bach, Lisbeth
; APPLICANT: No. 6100053revang, Iben
; APPLICANT: Haikier, Torben
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Schaefer, Thomas
; APPLICANT: Andersen, Jens
; TITLE OF INVENTION: Microbial Transglutaminases, Their
;     PRODUCTION AND USE
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 61000530 No. 6100053disk of No. 6100053th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793.426A
; FILING DATE: 25-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol E.
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 4211.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 331 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-793-426A-2

Query Match          80.9%; Score 1465; DB 3; Length 331;
Best Local Similarity 79.3%; Pred. No. 6.5e-137;
Matches 261; Conservative 34; Mismatches 34; Indels 0; Gaps 61

Qy   2 SDDRVTTPAEPLDRMPDPYPSPYGRAETVVNNYIRKWOOVYSHRGKRQQWTEQEWLS 61
      |||

```

```
QY 2 SDDVTPPAELDRMDPPKPSYGRAETVYNNYIRKQOVSHRDGRKQOMTEEOEWLS 61
Db 2 ADEKVTTPPAELNRPDPAVYRAYGRATVVYNNYIRKQOVSHRDGIQOOMEEOEKUS 61
QY 62 YGCVGTWVNSGOPYNRLAFASFEDRFKNELKNGRPSRGETRAFEGRVAKESFDEK 121
Db 62 YGCVGTWVNSGOPYNKLAFASFEDENKYKSDLENSRPNETQAEFEGRIVKDSPEK 121
QY 122 GFQARVAVSMRLENADHESAYLDNLKELANGNDALRNEDARSPYSALRNPSPK 181
Db 122 GFRRADVAVSMKALDSAHDEGTIDNLKTELANKDRLYEDGRNSFYLSALRNPSPK 181
QY 182 ERGGNHDPKMKAVIYSHFWGQDRSSADKRYKGDPAFPAEPTGLVDMKSRDRNP 241
Db 182 ERDGGNYDPSKMKAVIYSHFWGQDRSSADKRYKGDPAFPAEPTGLVDMKSRDRNP 241
QY 242 RSTPSCEGVNFDYGFAGQATADAKTWTWTHGNHYHAPNGSLGAMHYESKFRNWSG 301
Db 242 RSPAQGESWVNFYGFAGQATESDADKTWTWTHANHYHAPNGSLGPMNVYESKFRNWSAG 301
QY 302 YSDFDRGAYVITPIKSWNTAPDKVQGW 330
Db 302 YADFDRGTIVITPIKSWNTAPAEVQGW 330

RESULT 8
US-08-793-426A-7
; Sequence 7, Application US/08793426A
; Patent No. 6100053
; GENERAL INFORMATION:
; APPLICANT: Bech, Lisbeth
; APPLICANT: No. 6100053revang, Iben
; APPLICANT: Halkier, Torben
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Schafer, Thomas
; APPLICANT: Andersen, Jens
; TITLE OF INVENTION: Microbial Transglutaminases, Their
; TITLE OF INVENTION: Production And Use
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6100053o No. 6100053disk of No. 6100053th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,426A
; FILING DATE: 25-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol E.
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 4211.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
```

```
; FRAGMENT TYPE: internal
US-08-793-426A-7

Query Match 32.6%; Score 591; DB 3; Length 126;
Best Local Similarity 82.5%; Pred. No. 4.6e-51;
Matches 104; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 165 DARSFYSALRNTPEFKERNNGNDPDSRMKAVIYSHFWGQDRSSADKRYKGDPAFR 224
Db 1 DSRSFYSALRNTPEFKERDGNYPDSRMKAVIYSHFWGQDRSGSDKRYKGDPAFR 60
QY 225 PAPGGLVDMKSRDRNPSPSPGEGFVNFYGFAGQATADAKTWTWTHGNHYHAPNGS 284
Db 61 PDQGTGLVDMKSRDRNPSPSPAKPCEFFVNFYGFAGQAEADAKTWTWTHANHYHAPBG 120
QY 285 LGAMHV 290
Db 121 MGPMBV 126

RESULT 9
US-09-294-565-7
; Sequence 7, Application US/09294565
; Patent No. 6190879
; GENERAL INFORMATION:
; APPLICANT: Bech, Lisbeth
; APPLICANT: No. 6190879revang, Iben
; APPLICANT: Halkier, Torben
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Schafer, Thomas
; APPLICANT: Andersen, Jens
; TITLE OF INVENTION: Microbial Transglutaminases, Their
; TITLE OF INVENTION: Production And Use
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6190879o No. 6190879disk of No. 6190879th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/294,565
; FILING DATE: 19-APR-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4211.224-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 126 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
US-09-294-565-7

Query Match 32.6%; Score 591; DB 4; Length 126;
```

Best Local Similarity 82.5%; Pred. No. 4.6e-51;
Matches 104; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 165 DARSFYALRNTPSKERNGNHDSRKNKAVIYSHFWSGQDRSSADKRYGDPDAFR 224

Db 1 DSRSFYALRNTPSKERDGGNYDSKMKAVYSHFWSGQDRSSADKRYGDPDAFR 60

QY 225 PAPTGLVDMSRDNIPRSPGEGFVNFYDYGWFGAQTADADKTVWTHGNHYHAPNGS 284

Db 61 PDGPTGLVDMSRDNIPRSPAKEGEFVNFYDYGWFGAQTADKTVWTHANHYHAPBGG 120

QY 285 LGAMHV 290

Db 121 MGPMBV 126

RESULT 10

US-08-793-426A-8
; Sequence 8, Application US/08793426A
; Patent No. 6100053
; GENERAL INFORMATION:
; APPLICANT: Bech, Lisbeth
; APPLICANT: No. 6100053revang, Iben
; APPLICANT: Halkier, Torben
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Schafer, Thomas
; APPLICANT: Andersen, Jens
; TITLE OF INVENTION: Microbial Transglutaminases, Their
; TITLE OF INVENTION: Production And Use
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6100053o No. 6100053disk of No. 6100053th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatetle
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,426A
; FILING DATE: 25-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol E.
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 4211.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: C-terminal
; US-08-793-426A-8

Query Match 11.5%; Score 208; DB 3; Length 40;

Best Local Similarity 90.0%; Pred. No. 6.6e-14;

Matches 36; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 292 ESKFRNWSGYSDFRGAYVITFIPKSWNTAPDKVKQGW 331

Db 1 ESKFRNWSGYSDFRGAYVITFIPKSWNTAPDKVKQGW 40

US-09-884-948-1.rai

RESULT 11

US-09-294-565-8
; Sequence 8, Application US/09294565
; Patent No. 6190879
; GENERAL INFORMATION:
; APPLICANT: Bech, Lisbeth
; APPLICANT: No. 6190879revang, Iben
; APPLICANT: Halkier, Torben
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Schafer, Thomas
; APPLICANT: Andersen, Jens
; TITLE OF INVENTION: Microbial Transglutaminases, Their
; TITLE OF INVENTION: Production And Use
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6190879o No. 6190879disk of No. 6190879th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/294,565
; FILING DATE: 19-APR-1999
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4211.224-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 40 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: C-terminal
; US-09-294-565-8

Query Match 11.5%; Score 208; DB 4; Length 40;

Best Local Similarity 90.0%; Pred. No. 6.6e-14;

Matches 36; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 292 ESKFRNWSGYSDFRGAYVITFIPKSWNTAPDKVKQGW 331

Db 1 ESKFRNWSGYSDFRGAYVITFIPKSWNTAPDKVKQGW 40

US-09-884-948-1.rai

RESULT 12

US-08-793-426A-6
; Sequence 6, Application US/08793426A
; Patent No. 6100053
; GENERAL INFORMATION:
; APPLICANT: Bech, Lisbeth
; APPLICANT: No. 6100053revang, Iben
; APPLICANT: Halkier, Torben
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Schafer, Thomas

Query Match 11.5%; Score 208; DB 3; Length 40;

Best Local Similarity 90.0%; Pred. No. 6.6e-14;

Matches 36; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 292 ESKFRNWSGYSDFRGAYVITFIPKSWNTAPDKVKQGW 331

Db 1 ESKFRNWSGYSDFRGAYVITFIPKSWNTAPDKVKQGW 40

US-08-793-426A-6

```
; APPLICANT: Andersen, Jens
; TITLE OF INVENTION: Microbial Transglutaminases, Their
; TITLE OF INVENTION: Production And Use
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 61000530 No. 6100053disk of No. 6100053th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,426A
; FILING DATE: 25-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rozek, Carol E.
; REGISTRATION NUMBER: 36,993
; REFERENCE/DOCKET NUMBER: 4211.204-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; US-08-793-426A-6

Query Match 10.4%; Score 189; DB 3; Length 52;
Best Local Similarity 70.8%; Pred. No. 7.4e-12;
Matches 36; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 106 AEFGRVAKESFDEKGFQRAREVASYMNRALENAHDESAYDLNKKELAN 156
|||||:| | :|:|||||:|:|:| | | | | | | | | | | | | | | |
Db 1 AEFGRVAKGXFDAFGFKRAREVASYMNRKALDSAHDEGTIIDHLKTELAN 51

RESULT 13
US-09-294-565-6
; Sequence 6, Application US/09294565
; Patent No. 6190879
; GENERAL INFORMATION:
; APPLICANT: Bech, Lisbeth
; APPLICANT: No. 6190879revang, Iben
; APPLICANT: Haikier, Torben
; APPLICANT: Rasmussen, Grethe
; APPLICANT: Schafer, Thomas
; APPLICANT: Andersen, Jens
; TITLE OF INVENTION: Microbial Transglutaminases, Their
; TITLE OF INVENTION: Production And Use
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6190879o No. 6190879disk of No. 6190879th America, Inc.
; STREET: 405 Lexington Avenue
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/294,565
; FILING DATE: 19-APR-1999
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Green, Reza
; REGISTRATION NUMBER: 38,475
; REFERENCE/DOCKET NUMBER: 4211.224-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; FRAGMENT TYPE: internal
; US-09-294-565-6

Query Match 10.4%; Score 189; DB 4; Length 52;
Best Local Similarity 70.6%; Pred. No. 7.4e-12;
Matches 36; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 106 AEFGRVAKESFDEKGFQRAREVASYMNRALENAHDESAYDLNKKELAN 156
|||||:| | :|:|||||:|:|:| | | | | | | | | | | | | | | |
Db 1 AEFGRVAKGXFDAFGFKRAREVASYMNRKALDSAHDEGTIIDHLKTELAN 51

RESULT 14
US-08-558-135-5
; Sequence 5, Application US/08558135
; Patent No. 6090631
; GENERAL INFORMATION:
; APPLICANT: Catterall, William A.
; APPLICANT: Sheng, Zu-Hang
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING
; TITLE OF INVENTION: FOR PRESYNAPTIC CALCIUM CHANNEL BLOCKERS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/558,135
; FILING DATE: 13-NOV-1995
; CLASSIFICATION: 436
; ATTORNEY/AGENT INFORMATION:
; NAME: Sharkey, Richard G.
; REGISTRATION NUMBER: 32,629
; REFERENCE/DOCKET NUMBER: 920010.602C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 amino acids
; TYPE: amino acid
; STRANDEDNESS:
```

QY 98 -----RPRSGETRAFFEGRAVE---SPDEEKGFQR-----AREVASVMNRALE 138
Db 145 IISLDYFVHGGPNGAHVCMVFE--VLGENLLSLIQSYGHRGVPVGVKQIAYQLLIAD 202
QY 139 NAHDESAYL-DNLKKE---LANGNDALRNEDARSPFYCALRNTPTSPFKERNGNHDPGRMK 194
Db 203 YLHRECGIHTDLKPNVLCIDQDALQHEAPATTSSPTSSNTSSSKTRNTGY---TAK 259
QY 195 AVIYSKHFWSGDRSSADKRYGDPDAFRPAPGTGLV----- 232
Db 260 APIIKR-----GQSDVNSAQERKTKAKNPTKNSKPAGQVIPSSTSTLSRFPSPLEGAYSE 315
QY 233 ----DMSDRNIPRSPSPGEGFVNFYGNFQAOTE-----ADADKTWTHGN----- 276
Db 316 ISLRDSQKHNSHPSPSSGDNLSILD-GVNGSOEPVPKITVTKIADLGNACWTRKHFTND 374
QY 277 ----HYHAPNGSLGAMHVVYESKERNWSEGYSDFD--RGAYVITEIPKSWNT 321
Db 375 VOTROYRSPVILGCR--WGASADCWSPACILFELLTGDL--FDPNGNS 421

Search completed: September 27, 2002, 12:38:29
Job time: 22 sec

QY 99 PRSGETRAFFEGRAVESDEEKGFORAREVASVMNRALENAHDESAYLDNLKKELANGN 158
Db 107 PQENRNNTKSVABFTVDORLQOQRAEDFLKQARHHRADPSAH----- 154
QY 159 DALRNEDARSPFYSA-----LRNTPSFE-----RNGGNHDPGRMKAVIYSKHFWGGDR 208
Db 155 -AAAGLDARPWAGSOEAELSRGPGYGRSDHQAREGGLEPPG-----FWEGE-- 201
QY 209 SSSADKRYGDPDAFRPAPGTGLVDSNRDN-IPRSPTSPGE 249
Db 202 --AERKAGDPRH-RHAHQVGGSGGSRSPRIGTADGE 239

RESULT 15
US-08-264-002-7
; Sequence 7, Application US/08264002
; Patent No. 5559019
; GENERAL INFORMATION:
; APPLICANT: GUI, JIAN-FANG
; APPLICANT: FU, XIANG-DONG
; TITLE OF INVENTION: NOVEL PROTEIN SERINE KINASE, SRPK1
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1890 Century Park East, Fifth Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/264,002
; APPLICATION NUMBER: US/08/264,002
; FILING DATE: 22-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: TUMARKIN PH.D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3590
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-264-002-7

Query Match 5.0%; Score 91; DB 1; Length 544;
Best Local Similarity 20.7%; Pred. No. 1.2;
Matches 85; Conservative 48; Mismatches 148; Indels 130; Gaps 18;
QY 27 AETVNNVIRKWOQVYSHRDGR-----KQMTREQRLSYGCGVTVWNS----- 72
Db 25 SEKVITGKNVKNVSEVDGKSMVEKVTHEENAEHYGGHPYVIGEEFHHRYVVE 84
QY 73 -----GOYPTNRLAF-----ASFDEDRFKNELKNG----- 97
Db 85 PKLGGHFSITWLAYDRAAKRRVALKVVRSAEHYRETISIDEIRLQKIREGDEKHLGKKH 144

TOPOLOGY: linear
US-08-558-135-5

Query Match 5.5%; Score 99; DB 3; Length 315;
Best Local Similarity 26.5%; Pred. No. 0.088;
Matches 43; Conservative 18; Mismatches 61; Indels 40; Gaps 7;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2002, 12:38:07 ; Search time 19.4 Seconds
(without alignments)
1639.462 Million cell updates/sec

Title: US-09-884-948-1
Perfect score: 1811
Sequence: 1 DSDDRVTTPAEPLDRMPDPV.....ITPIPKSWNTAPDKVKQWGP 331
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_71.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1811	100.0	406	JC2089	protein-glutamine
2	1802	99.5	332	JC7310	protein-glutamine
3	114.5	6.3	1217	S57114	sericin1B - silkwo
4	106.5	5.9	918	E82486	proteinase VCA0223
5	105.5	5.8	1888	T14273	zinc finger protei
6	105	5.8	325	T38308	hypothetical prote
7	105	5.8	920	T52426	dynammin-like prote
8	105	5.8	1092	H96619	protein T30E16.17
9	104	5.7	575	S03745	beta-amylase (EC 3
10	102	5.6	219	B90045	hypothetical prote
11	102	5.6	504	T31784	hypothetical prote
12	101.5	5.6	283	AH3457	acriflavin resista
13	100.5	5.5	926	B37271	A-alpha Y 3 protei
14	100	5.5	517	A39038	1-caldesmon, nonmu
15	99	5.5	2273	I46477	calcium channel BI
16	99	5.5	2424	I46480	calcium channel BI
17	98.5	5.4	1507	B47328	natural killer cel
18	98.5	5.4	2924	T18378	variant-specific s
19	98	5.4	350	S25525	outer membrane por
20	98	5.4	547	A55575	puff-specific nucl
21	98	5.4	1403	A47328	natural killer cel
22	98	5.4	3938	T42761	Bassoon protein -
23	97.5	5.4	625	A34615	autotaggrin - rat
24	97	5.4	572	S18732	proantigen, 64k -
25	97	5.4	591	D89783	RGD-containing lip
26	97	5.4	1097	T31504	hypothetical prote
27	96.5	5.3	548	T28910	hypothetical prote
28	96	5.3	350	AF0543	outer membrane por
29	96	5.3	992	B86237	protein F14N23.17

30 96 5.3 1344 2 T42637 hypothetical prote
31 96 5.3 1883 2 T13944 chromodomain-helic
32 95 5.2 608 2 T32923 hypothetical prote
33 95 5.2 1233 1 A44400 myosin heavy chain
34 95 5.2 1313 2 A48467 myosin heavy chain
35 94.5 5.2 483 2 T19720 hypothetical prote
36 94 5.2 462 2 H61445 troponin T - fruit
37 93.5 5.2 396 2 S13251 hypothetical prote
38 93.5 5.2 1261 2 G83162 retinoblastoma bin
39 93.5 5.2 1722 1 I78879 respiratory nitrat
40 93.5 5.2 1993 2 AF1450 probable peptidogl
41 93.5 5.2 2441 2 D71623 erythrocyte membra
42 93 5.1 632 2 E75057 peptidase PAB1418
43 92.5 5.1 506 2 G90000 hypothetical prote
44 92.5 5.1 570 1 S50933 myb-related protei
45 92.5 5.1 707 1 S69781 outer membrane pro

ALIGNMENTS

RESULT 1
JC2089
protein-glutamine gamma-glutamyltransferase (EC 2.3.2.13) precursor - Streptovorticil
N: Alternate names: glutamyl-peptide--amine gamma-glutamyltransferase; transglutamin
C: Species: Streptovorticillum sp.
C: Date: 14-Jul-1994 #sequence revision 14-Jul-1994 #text_change 07-May-1999
C: Accession: JC2089; JC2090; A46730
R: Washizu, K.; Ando, K.; Koikeda, S.; Hirose, S.; Matsuura, A.; Takagi, H.; Motoki, M
Biosci. Biotechnol. Biochem. 58, 82-87, 1994
A: Title: Molecular cloning of the gene for microbial transglutaminase from Streptover
A: Reference number: JC2089; MUID: 94162748
A: Accession: JC2089
A: Molecule type: DNA
A: Residues: 1-406 <WAS>
A: Experimental source: strain S-8112
R: Takehana, S.; Washizu, K.; Ando, K.; Koikeda, S.; Takeuchi, K.; Matsui, H.; Motoki,
Biosci. Biotechnol. Biochem. 58, 88-94, 1994
A: Title: Chemical synthesis of the gene for microbial transglutaminase from Streptove
A: Reference number: JC2090; MUID: 94162749
A: Accession: JC2090
A: Molecule type: DNA
A: Residues: 76-406 <TAK>
R: Kanaaji, T.; Ozaki, H.; Takao, T.; Kawajiri, H.; Ide, H.; Motoki, M.; Shimonishi, Y.
J. Biol. Chem. 268, 11565-11572, 1993
A: Title: Primary structure of microbial transglutaminase from Streptovorticillum sp.
A: Reference number: A46730; MUID: 93280110
A: Accession: A46730
A: Status: preliminary
A: Molecule type: protein
A: Residues: 76-406 <KAN>
A: Experimental source: s-8112
A: Note: sequence extracted from NCBI backbone (NCBIP:133222)
C: Comment: This enzyme catalyzes an acyl transfer reaction between a gamma-carboxyami
lar crosslinking of certain proteins by gamma-glutamyl-epsilon-lysine side chain brid
C: Superfamily: protein-glutamine gamma-glutamyltransferase
C: Keywords: aminocyclotransferase; calcium; coagulation; heterotetramer; homodimer
F: 1-18/Domain: signal sequence #status predicted <Sig>
F: 19-75/Domain: propeptide #status predicted <PRO>
F: 76-406/Product: protein-glutamine gamma-glutamyltransferase #status predicted <MAT>

Query Match 100.0% Score 1811; DB 2; Length 406;
Best Local Similarity 100.0%; Pred. No. 1.9e-128;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 DSDDRVTTPAEPLDRMPDPYRPSYGRAETVYNNYIRKQWQYVSHRDGRKQOMTEQREWL 60
Db 76 DSDDRVTTPAEPLDRMPDPYRPSYGRAETVYNNYIRKQWQYVSHRDGRKQOMTEQREWL 135
QY 61 SYGCVGVTVWNSGQYPTNRLAFASDFEDFRFKNELKNGRPSRGETRAEFEGYVAKESDEE 120
Db 136 SYGCVGVTVWNSGQYPTNRLAFASDFEDFRFKNELKNGRPSRGETRAEFEGYVAKESFDEE 195

QY	121	KGFORAREVASYMNRALENAHDSEAYLDNLKKELANGNDALRNEDARSPFFYSALRNTPSF	180
DB	196	KGFORAREVASYMNRALENAHDSEAYLDNLKKELANGNDALRNEDARSPFFYSALRNTPSF	255
QY	181	KERNNGNHDPSRMKAVIYSKHFWSQDSSADKKYGDPAFRPAPGTGLVDMRSRDN	240
DB	256	KERNNGNHDPSRMKAVIYSKHFWSQDSSADKKYGDPAFRPAPGTGLVDMRSRDN	315
QY	241	PRSPITSPGEGFYNDYGVFGAQTADADKTVMTHGNHYPAPNGSLGAMHVYESKFRNWSE	300
DB	316	PRSPITSPGEGFYNDYGVFGAQTADADKTVMTHGNHYPAPNGSLGAMHVYESKFRNWSE	375
QY	301	GYSDFDRGAYVITFIPKSNWNTAPDKVKOGWP	331
DB	376	GYSDFDRGAYVITFIPKSNWNTAPDKVKOGWP	406
RESULT	2		
JC7310			
protein-glutamine gamma-glutamyltransferase (EC 2.3.2.13) - Escherichia coli			
N:Alternate names: microbial transglutaminase			
C:Species: Escherichia coli			
C:Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 17-Nov-2000			
C:Accession: JC7310			
R:Yokoyama, K.; Nakamura, N.; Seguro, K.; Kubota, K.			
Biosci. Biotechnol. Biochem. 64, 1263-1270, 2000			
A:Title: Overproduction of microbial transglutaminase in Escherichia coli, in vitro refd			
A:Reference number: JC7310			
A:Accession: JC7310			
A:Molecule type: DNA			
A:Residues: 1-332 <YOK>			
A:Experimental source: strain JM109			
A:Note: Cross-reference			
C:Comment: This enzyme belongs to a family of enzymes that catalyzes acyl transfer between			
in the formation of epsilon-(gamma-glutamyl) lysine cross-linkages. This enzyme is invol			
C:Genetics:			
A:Gene: mtg			
C:Superfamily: protein-glutamine gamma-glutamyltransferase			
C:Keywords: aminocyltransferase			

Query Match	99.5%;	Score 1802;	DB 2;	Length 332;
Best Local Similarity	99.7%;	Pred. No. 6.9e-128;		
Matches 330;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	DSDDRVTTPAEPLDRMPDPRPSYGRAETVYVNYIRKWOQVYSHRDGRKQOQTEQRWL	60	
Db	2	DSDDRVTTPAEPLDRMPDPRPSYGRAETVYVNYIRKWOQVYSHRDGRKQOQTEQRWL	61	
QY	61	SYGCGVGTWVWSGQPTNRLAFASDFDEFKNELNGRPRSGETRAETEGRAVAKESFDEE	120	
Db	62	SYGCGVGTWVWSGQPTNRLAFASDFDEFKNELNGRPRSGETRAETEGRAVAKESFDEE	121	
QY	121	KGQRAREYASVMNRALENAHDESAYLDNLKKELANGNDALNEDARSGFFYSALRNTSPF	180	
Db	122	KGQRAREYASVMNRALENAHDESAYLDNLKKELANGNDALNEDARSGFFYSALRNTSPF	181	
QY	181	KERNGNHNDDPRMKAVIYSKHFWSGODRSSADKKYGDPAFRAPGTGLVDMGRDNI	240	
Db	182	KERNGNHNDDPRMKAVIYSKHFWSGODRSSADKKYGDPAFRAPGTGLVDMGRDNI	241	
QY	241	PRSPSPGCGGFVNDYCGWFAGQATEADAKTVTHGHNHAPNGSLGAMHVYESKERNWSE	300	
Db	242	PRSPSPGCGGFVNDYCGWFAGQATEADAKTVTHGHNHAPNGSLGAMHVYESKERNWSE	301	
QY	301	GYSDFDORGAYVITFIKPSWNTAPDKVKQGW	331	
Db	302	GYSDFDORGAYVITFIKPSWNTAPDKVKQGW	332	

RESULT
S52714
3

sericinIb - silkworm
C:Species: Bombyx mori (silkworm)
C:Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 09-Sep-1997
C:Accession: S52714
R:Garel, A.A.; Deleage, G.G.; Prudhomme, J.J.
submitted to the EMBL Data Library, March 1995
A:Description: Structure and organisation of the Bombyx mori sericin I gene and of th
A:Reference number: S52714
A:Accession: S52714
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1217 <GAR>
A:Cross-references: EMBL:Z48802; NID:g755699; PID:g755700

Query Match	6.38;	Score 114.5;	DB 2;	Length 1217;
Best Local Similarity	20.4%;	Pred. No. 1.8;		
Matches 60;	Conservative 47;	Mismatches 136;	Indels 51;	Gaps 11;
Qy 33	NYRKWQGVYSHRGDKQOMTEQR-	--EWLSYGVGVYTWNSGQ-----	YPTNR 79	
Db 124	NYVSDGAVASSDARDENRSQAQNAQANWADSYGYVSADRSASASRRRCQANYYSKD 183			
Qy 80	LAFASFDEDRFKNELKNRPRSGETRAFFEGEVAKESDEEKGQQRAREVASVNNRALEN 139			
Db 184	ITAASKDSDRADSSRRSN-----	AYNFRDSD-----	GSESAGLSDRSASS 223	
Qy 140	AHDESAYLDNLKELANGENDALRNEDARSPYSALRNTPFPERKNGGNHDP2-----	RMK 194		
Db 224	SKNDNVFYRTKDSI--GGQAKSSRSSHSQESDAYNSSPDGYNAGTRDSTSNKKKAS 281			
Qy 195	AVIYS--KHFWSGGDRSSADKRKYGDPAFRPAGTGLVMSDRDNTPRSP7SGEGFV 252			
Db 282	STIYADKQIIRAANDRSSKOLKQSSAIGSGP-KGT5V--SSKDRQYNDKRRSKSDAYV 338			
Qy 253	NFDYGWFG--AQTADADKTYVTHGNHYHAPNGSLGAMHVFYKESFRN5GEYSD 304			
Db 339	GRD-----CTIVAYSNKDSKTSRQSTNTNADONSVERSDSAADQTSKSYDRGYSD 388			

RESULT

B24245
 proteinase VCA0223 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)
 C:Species: *Vibrio cholerae*
 C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
 C:Accession: E82486
 R:Heideberg, J.F.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.
 C:Johnson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers
 L., R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
 Nature 406, 477-483, 2000
 A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.
 A:Reference number: A82035; MUID:20406833
 A:Accession: E82486
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-918 <HEI>
 A:Cross-references: GB:AE004362; GB:AE003853; NID:q9657611; PIDN:AAF96135.1; GSPDB:CN
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor
 C:Genetics:
 A:Gene: VCA0223
 A:Map position: 2
 C:Superfamily: *Bacillus thuringiensis* immune inhibitor A

Query Match	5.94;	Score 106.5;	DB 2;	Length 918;
Best Local Similarity	21.25;	Pred. No. 5.1;		
Matches 65;	Conservative 52;	Mismatches 113;	Indels 77;	Gaps 16;
QY	42	YSHR-DGRRKQMTPEQREWLSCYGVGVTVWNSGQYPTNRL-AFASFDED-RFKNELKNGR	98	
		: : :		
DB	446	YSNRGDDLLKNSRP-----LITPAGSQATLRFKAWFOIEKDYARVYLNGK	493	
QY	99	PRSGTGRA-----EFGKRVAKESDFDEKGFQKAR-EVAVSYNNRALENAHD-----ESA	145	

Db 494 PIAGNITMDDPKSLPAISQSGWYDAQFDLSAWAGQIVELAFDLYLTGGLAMEGL 553
QY 146 YLDNLKELANGDALRNDARSPPFSALENTPSK--PRNGNHDPSPMKAVIYSKHPW 203
Db 554 YVDLLEVDGNOTLLDIAEGTSSF-----AFQFTKNGGFHE-----ANHY 596
QY 204 SGDRSSS-----ADKRYGDPDAPRPAPGTGLVDMRDRIIPRSPTSPGSGFVNFYD 257
Db 597 LLQWRSHNDVQGLANKRFGQLMSEPLLWYVDESADNWVGK--HPGE-----G 647
QY 258 WFGAQTAEADAKTVWTHGHNYHAPNSGLGAMHVESKFRNWSGSDFRGAYVTFIPK 317
Db 648 WLGV-VDAONALVWSKTG-----EVAQTRFQVRDATTSLFDQAP--LKLVA 692
QY 318 SWNTAPD 324
Db 693 DGNLTED 699
RESULT 5
T14273
zinc finger protein 106 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14273
R:Zuberi, A.R.; Christianson, G.J.; Mendoza, L.M.; Shastri, N.; Roopenian, D.C.
Immunity 9, 687-698, 1998
A:Title: Positional cloning and molecular characterization of an immunodominant cytotoxic
A:Reference number: Z17953; MUID:99060924
C:Accession: T14273
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1888 <ZUB>
A:Cross-references: EMBL:AF060246; NID:93372656; PID:93372657; PIDN:AAD04329.1
C:Genetics:
A:Note: zip106
Query Match 5.8%; Score 105.5; DB 2; Length 1888;
Best Local Similarity 19.6%; Pred. No. 15;
Matches 68; Conservative 37; Mismatches 101; Indels 141; Gaps 17;
QY 8 PPAELDRMPDPYRPSYGRAETVYNNYIRKQQVYSHRDGRKQOMTEQREMLSYGCVGV 67
Db 109 PPSNSQVNSDDRQPPQWRREDPIYQDRESYQPPRHKG-----PPQRDW----- 154
QY 68 TWNSGQYPTNRLAFASFDDEKFKELKN-GRPR-----SGETRA----- 106
Db 155 KWEKDG-----FNSTRKNSPFLRLNSGGPRGSSVWHKGATRGSSTWFLNHSNGGG 206
QY 107 -----EFEGRVAKESFDEE--KGFORAREVASVMNRALENHADESAYLDNLKEL 154
Db 207 WHSNNGMVDNNGTGRNSSWHSEGTGGF-----PSWHMNN----- 242
QY 155 ANGDALRNDARSPPFYSALRNT-----PSPKERNCGNHDPSPK-----K 194
Db 243 SNGN-----WKSSYRSTNSNNGYNGPGDKFOQGRNRNPNYQMEDTKMWNKSNK 291
QY 195 AVIYSKH--FWSGQDRSSSADKRYGDPDAPRPAPGTGLVDMRDRIIPRSPTSPGSGEV 252
Db 292 PSKYSQERCKWQRDRKAAKYR--SPPEGY-----ASDTFPSEGLL 331
QY 253 NFDYWGFAQTAEADAKTVWTHGHNYHAPNSGLGAMHVESKFRNWS 299
Db 332 EFNP-----EQRESQTKQTDATAKINGKNGTKA-----RDKFPRWT 369
RESULT 6
T38308
hypothetical protein SPAC23H3.15c - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 20-Oct-2000

C:Accession: T38308; T38381
R:Skellton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.; Wood, V.
submitted to the EMBL Data Library, September 1997
A:Reference number: Z21785
A:Accession: T38308
A:Molecule type: DNA
A:Residues: 1-325 <SKE>
A:Cross-references: EMBL:Z99163; PIDN:CAB16245.1; GSPDB:GN00066; SPDB:SPAC23H3.15c
A:Experimental source: strain 972h-; cosmid c23H3
R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, August 1997
A:Reference number: Z21737
A:Accession: T38381
A:Molecule type: DNA
A:Residues: 1-87 <DEV>
A:Cross-references: EMBL:Z98944; PIDN:CAB11598.1; GSPDB:GN00066; SPDB:SPAC25H1.01c
A:Experimental source: strain 972h-; cosmid c25H1
C:Genetics:
A:Gene: SPAC23H3.15c; SPDB:SPAC25H1.01c
A:Map position: 1

Query Match 5.8%; Score 105; DB 2; Length 325;
Best Local Similarity 23.0%; Pred. No. 1.7;
Matches 64; Conservative 33; Mismatches 85; Indels 96; Gaps 16;
QY 71 NSGQYPTNRLAFASFDDEKFKN-----ELKNRPRSGETRAEFEGRVAKESFDEE 120
Db 32 SSSGAPQRNFDTSYTSEIPNSRAANDMTDGGDPYAGMTS-----DTK 79
QY 121 KGF-----ORAREVASVM--NRALENHADESAYLDNLKELANGNDAL----RNEDARSPE 170
Db 80 KGFNSVESRKKEQSDVRGDDTYSRRHDDSSYSN--KYSTGGNDSSYSGGRED----- 132
QY 171 YSALRNTSFERNCGNH--DPSRM-KAVIYSKHFWs-----GQDRSSSADKRR 216
Db 133 YST-----SGSYTTDPSTRTDSTASYGQSQYNGSRKTTGGDYGEDYSQSYPTDT 182
QY 217 YGDDPAFRPAPGTGLVDMRDRIIPRSPTSPGSGEVNFYDYGWFAQTAEADAKTVWTHGN 276
Db 183 YG-----SRQKATPSDVTGGGA----YDSSSGSHTHGSHGTEHRGS 222
QY 277 H--YHAPNGSGAMHVESKFRNWSGYS--DFDRGAY 310
Db 223 YGNDNTANKTRGAVS-----SAGSYGEGYKGTY 251
RESULT 7
T52426
dynamain-like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 24-Oct-2000
C:Accession: T52426
R:Kato, A.; Suzuki, M.; Kuwahara, A.; Ooe, H.; Higano-Inaba, K.; Kameda, Y.
Gene 239, 309-316, 1999
A:Title: Isolation and analysis of cDNA within a 300 kb Arabidopsis thaliana genomic
A:Reference number: Z25171
A:Accession: T52426
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-920 <KAT>
A:Cross-references: EMBL:AB028467; PIDN:BAA88111.1
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: CFI
A:Map position: 1

Query Match 5.8%; Score 105; DB 2; Length 920;
Best Local Similarity 24.6%; Pred. No. 6.6;
Matches 59; Conservative 27; Mismatches 104; Indels 50; Gaps 11;
QY 26 RAETVY--NNYIRKQQVYSHRDG-----RKQOMTEQREMLSYG 63

```

Db 684 KAESVMDKNEWINKLOKVIQARGQGVGSASMRQSLSEGSLDKMVRKVPDPEELRMSQE 743
QY 64 CVGVTWVNSGOYPTN-----RLAFASFDDEDFKNEKLNKGRPRSGETRAEF---EGRVAK 115
Db 744 VRGYVEAVLSLAANVPKAVVLCQVEKSKEDMLNQLYSSISAIGNERIESLIQEDQNVKR 803
QY 116 SFDEEKGFORAREVASVMNRALENHADE-----SAYLDNLKXELA---NGNDALRNEDAR 167
Db 804 RRDR---YKQSSLLSKLTRQL-SIHDRNRAAASWSNDSGTSPRTNGSS--GEDWM 857
QY 168 SPFYSALENTPTSPKERNNGNHDPSRMKAVIYKHFWSGGDRSSADKRYGDPDAFRPAP 227
Db 858 NAFNAAAGPDSLRYGSGGH--SRR-----YSDPAQNGEDSSGSGSSRRTPNRLPPAP 911

RESULT 8
H96619
protein T30E16.17 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H96619
R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso-
Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Luo, Z.A.; Lueros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: H96619
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1092 <STO>
A:Cross-references: GB:AE005173; NID:g8778745; PIDN:AAF79753.1; GSPDB:GN00141
C:Genetics:
A:Gene: T30E16.17
A:Map position: 1

Query Match 5.8%; Score 105; DB 2; Length 1092;
Best Local Similarity 24.6%; Pred. No. 8.2;
Matches 59; Conservative 27; Mismatches 104; Indels 50; Gaps 11;
QY 26 RAETVV--NNYIRKQOVYSHRDG-----RQQTDEOREWLSYG 63
Db 856 KAESVMDKNEWINKLOKVIQARGQGVGSASMRQSLSEGSLDKMVRKVPDPEELRMSQE 915
QY 64 CVGVTWVNSGOYPTN-----RLAFASFDDEDFKNEKLNKGRPRSGETRAEF---EGRVAK 115
Db 916 VRGYVEAVLSLAANVPKAVVLCQVEKSKEDMLNQLYSSISAIGNERIESLIQEDQNVKR 975
QY 116 SFDEEKGFORAREVASVMNRALENHADE-----SAYLDNLKXELA---NGNDALRNEDAR 167
Db 976 RRDR---YKQSSLLSKLTRQL-SIHDRNRAAASWSNDSGTSPRTNGSS--GEDWM 1029
QY 168 SPFYSALENTPTSPKERNNGNHDPSRMKAVIYKHFWSGGDRSSADKRYGDPDAFRPAP 227
Db 1030 NAFNAAAGPDSLRYGSGGH--SRR-----YSDPAQNGEDSSGSGSSRRTPNRLPPAP 1083

RESULT 9
S03745
beta-amylase (EC 3.2.1.2) precursor - Bacillus circulans
C:Species: Bacillus circulans
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S03745
R:Siggens, K.W.
Mol. Microbiol. 1, 86-91, 1987
A:Title: Molecular cloning and characterization of the beta-amylase gene from Bacillus d
```

```

A:Reference number: S03745; MUID:88260890
A:Accession: S03745
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-575 <SIG1>
A:Cross-references: GB:F00523; NID:g39407; PIDN:CAA68578.1; PID:g39408
C:Function:
A:Description: catalyzes the hydrolysis of 1,4-glycosidic linkages of starch, removin
C:Superfamily: beta-amylase
C:Keywords: glycosidase; hydrolase; polysaccharide degradation
F:1-36/Domain: signal sequence #status predicted <SIG>
F:37-575/Product: beta-amylase #status predicted <MAT>

Query Match 5.7%; Score 104; DB 1; Length 575;
Best Local Similarity 22.9%; Pred. No. 4.3;
Matches 76; Conservative 37; Mismatches 103; Indels 116; Gaps 19;
QY 24 YGRAETVVNNYIRKQOVYSHRDGKRCQOMTEEQREWLSYGCVGVTWV-----NSG 73
Db 80 WGVYESAGDNQF-DWSYKYTYADTVKQ-----AGLKWVPIIITHRCGGNVG 124
QY 74 Q-----YPTNRLAFASDEDFKNEKLNKGRPSG---ETRAEFGHVAKEPDEEKGFG 124
Db 125 DDCNIPLSFWSKSGADEMOKDE-----SGYVNESLSPFWSGVGKQ-YDE----- 171
QY 125 RAREVASVMNRALENHADESAYLDNLKKEANGDNLARNEDARSPFYSALENTPTSPKERN 184
Db 172 ---LYASP-----AQNFSAYKDMIPRIYLSGGP-----SGELRYPSYYPAA 209
QY 185 GGNHDPSPRMKAVIYKHFWSGGDRSSADKRYGDPDAFRPAPGTGLVDMSRDRNIPRSP 244
Db 210 GWSY-PARGKQVYIE---TAKSAFRMTATKYSLKXINAAGNTLTSMSQ-----ISP 260
QY 245 TSPGEGF-----VNFYDG-----WF-----GAQTEADADKTVWTH-----GNH 277
Db 261 PTDSGFGYTGGINITYKDFLSWYQSVLENHGLGVIGAAAHKDFPVGVIRIGAKISGIH 320
QY 278 YHAPNGSL--GAMHVVYESKERNWSEGYSDFD 307
Db 321 WQMNPNMPHSAEH-----AGGYDYDNR 343

RESULT 10
B90045
hypothetical protein [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B90045
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B90045
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-219 <KUR>
A:Cross-references: GB:BA000018; PID:g13702382; PIDN:BAB43523.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA2221

Query Match 5.6%; Score 102; DB 2; Length 219;
Best Local Similarity 20.8%; Pred. No. 1.7;
Matches 46; Conservative 42; Mismatches 73; Indels 60; Gaps 11;
QY 18 DYPFSPY-----GRAETVVNNYIRKW-----QQVYSHRDGKRCQOMTEEQRE-- 58
Db 5 DKYRSQYDQNNQNRHQSDASYSQQYAKGDPPEHPERYNGDYRREQLLEENK 64
```

QY 59 -----WLSYGVGVWNSQVPTNRLAFASFDDEDFKLNKNGRSPRSGETRAEFEGRY- 112
 Db 65 RRSKKWL-YIIIIALLIIVAFVTR--ALLNDSKVSNDPKVSNQYKKQVENQ-DGQIN 120
 QY 113 -----AKESFDEKGFQRAEVASVNNRALENAHDSA-----YLDNLKKELANG 157
 Db 121 QQVDNAKENI---KNQKTDIIKLNQIDNLKQEQNKADSKLTQFYQDOINK-LTEA 176
 QY 158 NDALRNEDARSPYSALRNTPS-----FKERNGG 186
 Db 177 NNALKNNASQKTESMLNDINTKFDISKLSLFDADNGG 217
 RESULT 11
 T31784
 hypothetical protein F13H6.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Jan-2000
 C:Accession: T31784
 R:Jones, K.; Wohlmann, P.
 submitted to the EMBL Data Library, July 1997
 A:Description: The sequence of C. elegans cosmid F13H6.
 A:Reference number: 221085
 A:Accession: T31784
 A:Status: preliminary;
 A:Molecule type: DNA
 A:Map position: 5
 A:Superfamily: cholinesterase; cholinesterase homology
 Query Match 5.6%; Score 102; DB 2; Length 504;
 Best Local Similarity 20.5%; Pred. No. 5.1;
 Matches 69; Conservative 44; Mismatches 144; Indels 80; Gaps 13;
 QY 28 ETVVNNYIRKQVYSHRDRKQKQTEQREWLSYQ---CVGVWNSQVPTNRLAFA 83
 Db 149 DSVNGNFGLDQTLA-----LKWQKHSSFGGDPNCVTVFGQAGGASTDLSLS 200
 QY 84 SFDEDFKLNKNGRSPRSGETRAEFEGRYAK-----ESFDEKGFQRAEVASV---NN 134
 Db 201 PHSKDLRFQFI---FISGAYCEPALRTSKQAKIFREAFKGTGDSITLLEWYKN 256
 QY 135 RALENAHDSAYLDNLKKELANGNDALRNEDAR-----SPFYALANTPSFKERNGGNH 189
 Db 257 QSSETLSD--LRKEAPKKQMTGVDEYGVIAAMNPEFSPADAGLALFFKGYGNDTAEN 314
 QY 190 PSRMKAVIYKSHFWSQDRSSSADKKK-----YGDPAFRPAPGTGLVDMRDRNIPRSP 244
 Db 315 PEEMHKLFEKVV-EGVDRSDSSAMKKRLCEAFG----- 348
 QY 245 TSPGEGFVNDYQWF-GAQTEDADKTVTHG-NHYHAPNGSL-----GAMHVIYSEK 294
 Db 349 -----LGFNLGVQSAKSSAKYKNDVFLYSFYHSDGFGMKDLPLFEASMHGTELR 401
 QY 295 FRNWKSEGYDFDRGAVYTFIPKSWNTAPDKVKQGW 331
 Db 402 YLLGEGFYSKFDATKEELEVLKTTILFSNFAKYNP 438
 RESULT 12
 AH3457
 acriflavin resistance protein E [imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 C:Accession: AH3457
 R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
 ; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess

Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit
 A:Reference number: AD3252; PMID:1175668
 A:Accession: AH3457
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-383 <KUR>
 A:Cross-references: GB:AE008917; PIDN:AAL52827.1; PID:gl7983667; GSPDB:GN00190
 A:Experimental source: strain 16M
 C:Genetics
 A:Gene: BMEI1646
 A:Map position: 1

Query Match 5.6%; Score 101.5; DB 2; Length 383;
 Best Local Similarity 26.1%; Pred. No. 3.9;
 Matches 29; Conservative 21; Mismatches 32; Indels 29; Gaps 3;
 QY 53 GCVGVTWNSQVPTNRLAFASFD-----DRFKNELNGRSPRSGETRAEFEGR 111
 Db 189 GIVGILPVNAGNYVTAQTSIARDSTVLIDWVPERFAIQKVGQPLTAESA-FPGE 247
 QY 112 VAKESFDEKGFQRAEVASVNNRALENAHDSAYLDNLKKELANGNDALR 162
 Db 248 IVKGRIN-----AVDNMLDEASRTLHVRAEVPNAEDRLR 281

RESULT 13
 B37271
 A-alpha Y 3 protein - bracket fungus (Schizophyllum commune)
 C:Species: Schizophyllum commune
 C:Date: 07-Feb-1992 #sequence_revision 07-Feb-1992 #text_change 05-Dec-1997
 C:Accession: B37271
 R:Ullrich, R.C.
 submitted to the Protein Sequence Database, October 1991
 A:Reference number: A37271
 A:Accession: B37271
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-926 <ULL>
 C:Superfamily: unassigned homeobox proteins; homeobox homology
 C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
 F:148-204/Domain: homeobox homology <HOX>

Query Match 5.5%; Score 100.5; DB 2; Length 926;
 Best Local Similarity 20.5%; Pred. No. 14;
 Matches 71; Conservative 36; Mismatches 122; Indels 117; Gaps 15;
 QY 9 PAEPLDRPDPYRSYGRAETVNNYIRKQVYSHRDRKQKQTEQREWLSYGCYGV 68
 Db 601 PAEPIVR-PDDFAFPVALAEKRAKRRARKEKKQAEKARK----- 640
 QY 69 WVNSQVPTNRLAFASFDDEDFKLNKNGRSPRSGETRAEFEGRYAKESFDEKGFQRA 128
 Db 641 -----DEKRAKRAKQAKKDRKEQRAGLPRR-SPSLDS-----SRASS 678
 QY 129 VASVNNRALENAHDSAYLDNLKKELANGNDALRNEDARSPFYALRNTPSKERNNGNH 188
 Db 679 VTSASATSRKSRKSRKPRDSSASSVAS-----ARTP---SLSTSS--RRSGTS 724
 QY 189 DPS--RMK---AVIYKSHFWSQDRSSAD---KRYKQDPDA----- 222
 Db 725 MPATPRMNEPLPVVASDNFVLGTDKDVTWTPELMAQLFGEDDASGLDBPMQSEGFSDML 784
 QY 223 -FRPAPGTGLVDMRDRNIPRSPSPGEGFYNFYGNWFG---AQTEADAKTVTHGNH 279
 Db 785 IFSSCNDGALGDMTADVKNPE-----LGDLSDTQLSFDMDNMTSSMDLS 828
 QY 280 A-PNGSLGAMHVIYSEKFRNWKSEGYDFDRGAVYTFIPKSWNTAPD 324
 Db 829 TOPAASFDSSETSSMDPNW-----LLPQCANTAPD 859

RESULT 14

A39038
l-caldesmon, nonmuscle - chicken
N:Alternate names: caldesmon, nonmuscle
C:Species: Gallus gallus (chicken)
C>Date: 31-Jul-1991 #sequence_revision 27-Jun-1994 #text_change 22-Jun-1999
C:Accession: A39038
R:Hayashi, K.; Fujio, Y.; Kato, I.; Sobue, K.
J. Biol. Chem. 266, 355-361, 1991
A:Title: Structural and functional relationships between h- and l-caldesmons.
A:Reference number: A39038; MUID:91093148
A:Accession: A39038
A:Molecule type: mRNA
A:Residues: 1-517 <RAY>
A:Cross-references: GB:M60620; GB:M38015; NID:g212242; PIDN:AAA48936.1; PID:g212243
A:Experimental source: brain
C:Comment: The binding of caldesmon to F-actin is modulated by calcium and calmodulin.
C:Comment: Two calmodulin molecules can bind to nonoverlapping domains of each caldesmon.
C:Superfamily: caldesmon
C:Keywords: actin binding; calmodulin binding; phosphoprotein
F:342,427,462/Binding site: phosphate (Ser) (covalent) #status predicted
F:433,456/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 5.5%; Score 100; DB 1; Length 517;

Best Local Similarity 18.7%; Pred. No. 7.4; Matches 60; Conservative 41; Mismatches 116; Indels 104; Gaps 12;

QY 28 ETVVNNYIR-KNQOVYSHRDGRKQKQMTQREWLISYGCVGVTWVNSGQYPTNRLAFASPD 86

Db 151 ETVTKSYQRNNRQ-DGEEGKKEKDSBEK-----PKEVPT-----E 188

QY 87 EDRFK-NELANGPRSGTFAEFGVAKESFDEERGFQARAREVASVMNRALENA----- 140

Db 189 ENQVKDKVKEAPKE-EMKSWD---RKGVPQQAQNGERELTPKLKSTENAFGRSN 244

QY 141 -----HDESYLNLKKELANGNDALRNEDARSPFYALRNTPSFKE 182

Db 245 LKGAANAAGSEKLKEKQQAAYVELDELKRRERKILLEEDEQKKQEAERKIREEE 304

QY 183 RSGNHDPSRMKAVIYKHPWSGDRSSADK-----RKYGDPD 221

Db 305 KRMKEIEIRRAEAEKROKVPEDGVSEKPKFCFSPKSGSLKTEERAEFLNKSQAKS 364

QY 222 AFRPAGTGLVDMSRDR-----NIPRSPTS-----PGEGFVNFYDGNFGAQT 263

Db 365 GMRPAHTTAVVSKIDSLRLEQYTSVVGNKAAPKAAPASDLPPVPAEGVRNI----- 415

QY 264 EADADKTWTHGNYHAPNGS 284

Db 416 -----KSMWEKGNVFSPPGT 431

RESULT 15

I46477
calcium channel BI-1 - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 17-Nov-2000
C:Accession: I46477; I46478
R:Mori, Y.; Friedrich, T.; Kim, M.S.; Mikami, A.; Nakai, J.; Ruth, P.; Bosse, E.; Hofmann
Nature 350, 398-402, 1991
A:Title: Primary structure and functional expression from complementary DNA of a brain c
A:Reference number: I46477; MUID:91187110
A:Accession: I46477
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2273 <MOR>
A:Cross-references: EMBL:X57476; NID:g1522; PIDN:CAA40714.1; PID:g1523
A:Accession: I46478
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1'1856, 'H',1858, 'K',1860-1862, 'SL',1865-1866, 'VIS',1870-1876, 'K',1878-1879,

A:Cross-references: EMBL:X57688; NID:g1524; PIDN:CAA40871.1; PID:g1525
C:Superfamily: voltage-dependent calcium channel protein alpha-1 chain

Query Match 5.5%; Score 99; DB 2; Length 2273;

Best Local Similarity 26.5%; Pred. No. 60; Matches 43; Conservative 18; Mismatches 61; Indels 40; Gaps 7;

QY 99 PRSGETRAEFEGRVAKESFDEERGFQARAREVASVMNRALENAHDSEAYLDNLKKELANGN 158

Db 828 PQENNNNTNKSRYAEPTVDQLGQQAEDFLRKQARHHRARDPESA----- 875

QY 159 DALRNEDARSPFYSA-----LRNTPSFKE-----RNGNHDPSRMKAVIYKHPWSGQDR 208

Db 876 -AAAGLDARRWAGSQEAELSREGPYGRESHQAREGGLPPG-----FWEGE-- 922

QY 209 SSSADKKRYGPDAPFRPAPGTGLVDMGRDRN-IPRSPTSPE 249

Db 923 --AERGKAGDPHR-RHAHRQGVGGSGSGSPRTGTADGE 960

Search completed: September 27, 2002, 12:38:57

Job time: 50 sec


```

QY 121 KGFQREAVSNRALENAHDESAVLDNLKKELANGNDALNEADARSPFYSALENTPSF 180
Db 166 KGFQREAVSNRALENAHDESAVLDNLKKELANGNDALNEADARSPFYSALENTPSF 225
QY 181 KERNGNHDPSRMKAVIYSKHFWSQDSSADKRYKGGPDPAFRPAGTGLVDMRSRDRNI 240
Db 226 KERNGNHDPSRMKAVIYSKHFWSQDSSADKRYKGGPDPAFRPAGTGLVDMRSRDRNI 285
QY 241 PRSPTSPGGGFVNFYDYGWFGAQTADADKTVTHGNHVNHPNGSLGAMHVYESKERNHSE 300
Db 286 PRSPTSPGGGFVNFYDYGWFGAQTADADKTVTHGNHVNHPNGSLGAMHVYESKERNHSE 345
QY 301 GYSDFDRGAYVITFIPKSNWNTAPDKVKGWP 331
Db 346 GYSDFDRGAYVITFIPKSNWNTAPDKVKGWP 376

RESULT 2
Q9RIS1 PRELIMINARY; PRT; 411 AA.
AC Q9RIS1;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE TRANSGUTAMINASE PRECURSOR (EC 2.3.2.13).
GN TGASE.
OS Streptomyces cinnamonensis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=53446;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CBS 683.68;
RX MEDLINE=98336622; PubMed=9672751;
RA Duran R., Junqua M., Schmitter J.M., Gancet C., Goulas P.;
RT "Purification, characterization, and gene cloning of transglutaminase
RT (Tgase) from Streptovorticillum cinnamonum CBS 683.68."
RL Biochimie 80:313-319(1998).
KW EMBL; Y08820; CAA70055.1;
KW Signal; Transferase; Acyltransferase.
FT SIGNAL 1 81 POTENTIAL.
FT CHAIN 82 411 POTENTIAL.
SQ SEQUENCE 411 AA; 46499 MW; 01A0CCA2EF4C388B CRC64;

```

```

Query Match 79.1%; Score 1432; DB 2; Length 411;
Best Local Similarity 79.7%; Pred. No. 3.4e-99;
Matches 263; Conservative 27; Mismatches 36; Indels 4; Gaps 3;

QY 2 SDRVTPPAELDRPDPYRPSYGRAETVNNYIRKWOQVYSHRDGRKQKQMTDQREWLS 61
Db 86 SDRVTPPAELDRPDPYRPSYGRAETVNNYIRKWOQVYSHRDGRKQKQMTDQREWLS 145
QY 62 YGCVGVTVNNSQYPTNRLAFASDFEDRFKNEKNGRPSRGTAEFGRVAKESFDEEK 121
Db 146 YGCVGVTVNNSQYPTNRLAFASDFEDRFKNEKNGRPSRGTAEFGRVAKESFDEEK 204
QY 122 QFORAREVSNRALENAHDESAVLDNLKKELANGNDALNEADARSPFYSALENTPSFK 181
Db 205 GFKRDVSNRALENAHDESAVLDNLKKELANGNDALNEADARSPFYSALENTPSFK 263
QY 182 ERNGNHDPSRMKAVIYSKHFWSQDSSADKRYKGGPDPAFRPAGTGLVDMRSRDRNI 241
Db 264 ERDGGNDPSRMKAVIYSKHFWSQDSSADKRYKGGPDPAFRPAGTGLVDMRSRDRNI 323
QY 242 RSPTSPGGGFVNFYDYGWFGAQTADADKTVTHGNHVNHPNGSLGAMHVYESKERNHSE 301
Db 324 RSPAKPGGWNFVNFYDYGWFGAQTADADKTVTHGNHVNHPNGSLGAMHVYESKERNHSE 383
QY 302 YSDFDRGAYVITFIPKSNWNTAPDKVKGWP 331
Db 384 YADF--GAYVITFIPKSNWNTAPAKVEQWP 411

```

```

RESULT 3
Q24441 PRELIMINARY; PRT; 549 AA.
ID Q24441;
AC Q24441;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE (CDNA2) PROTEIN 4.1 HOMOLOGUE (CORACLE) (FRAGMENT).
GN CORA OR CORACLE OR CG11949.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON R;
RX MEDLINE=94215495; PubMed=8162854;
RA Fehon R.G., Dawson I.A., Artavanis-Tsakonas S.;
RT "A Drosophila homologue of membrane-skeleton protein 4.1 is associated
RT with septate junctions and is encoded by the coracle gene."
RL Development 120:545-557(1994).
DR EMBL; L27468; AAA28742.1;
DR FlyBase; FBgn010434; cora.
KW Alternative splicing.
FT NON_TER 1
SQ SEQUENCE 549 AA; 59284 MW; A82055EF2BBA4874 CRC64;

Query Match 6.3%; Score 114.5; DB 5; Length 549;
Best Local Similarity 21.3%; Pred. No. 1.5;
Matches 71; Conservative 42; Mismatches 140; Indels 81; Gaps 15;

QY 12 PLDRMPDPYRPSYGRAETVNNYIRKWOQVYSHR-----DGRKQKQMT 54
Db 68 PVDRTPPKFNRTLSGARLTGRSDALALAEKVKARSTLDRGNADGDAHSRSPK 127
QY 55 EQREWLSYGCVGVTWNS--GQVTPNRLAFASD-----EDRKNEKNGRPSRGT 104
Db 128 NKREKSTGTASASSQSLEGGYETN-LEIEAIEAPFPVODADKEAKLRKKKEEKE 186
QY 105 RAIEFGRVAKESFDEEKFORAREVSNRALENAHDESAVLDNLKKELANGNDALRNE 164
Db 187 RKEREKRELEKKKAEEKAAALAGAAAGAAV-NGNDE--LNSNKSCKSSG----RRV 239
QY 165 DARSPPYSALRNTSPSKERNGNHDP--SRMKAVIYSKHFWSQDSSADKRYKGGPD 222
Db 240 DPNDPRFAGARTTVTHMTLTGIDPVTGRKS-----EYGDID- 278
QY 223 FRPAGTGLVDMRSRDRNIPRSPSPGEG-----FVNFDYGWFG--AQTADADKTVTHG 275
Db 279 ----PNTGDDIPA-----TAVTDPVTGKLILNVAOIDSFSGKQAOVTTTETVPIRQ 328
QY 276 NHVHAPNGSLGAMHVYESKERNHSEKNSGYSDFDRGA 309
Db 329 QFFD-----GVKHISKGLRRDSEGSDDDMTA 356

RESULT 4
Q9V8R8 PRELIMINARY; PRT; 889 AA.
ID Q9V8R8;
AC Q9V8R8;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CORA PROTEIN.
GN CORA OR CG11949.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

```


RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA April J.F., Aghayani A., An H.-J., Andrews-Pfankuch C., Miklos G.L.G.,
RA Bailly R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke K., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster A., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwac C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
DR EMBL: AEO03449; AAF46559.1; .
DR Flybase: FBgn0030182; CGI5311.
DR InterPro: IPR002106; AA_TRNA_ligase_II.
DR PROSITE: P500339; AA_TRNA_LIGASE_II-2; UNKNOWN_1.
SQ SEQUENCE 1161 AA; 130214 MW; DD01D5B408051D36 CRC64;

Query Match 6.2%; Score 112.5; DB 5; Length 1161;
Best Local Similarity 19.1%; Pred. No. 5.8; Mismatches 152; Indels 65; Gaps 13;
Matches 67; Conservative 66;

QY 8 PPAEPLDRMPDPYRPSYGRAETVNNYIRKQVYSHRDGRKQMQTEQREWLSCYGV 67
Db 180 POLEDREQVEHEENAFGR-QSFAYKKLNMHEQQSQQDKRGDDGDE----- 226
QY 68 TWVNSQYTNRLAFASFDEDFRKNELKNGRP---RSGETRAEFGR-----VAKESFDE 119
Db 227 VDLGGEYPPSQL-----FPAIEILNERKFNKPSAGSTGKSRPQLQKKDKQDMQEQQV 282
QY 120 EKGFORAREVASVMNRALNAHDESAYLDNLKKELANGNDALRNEDARSPPYSALRNTPS 179
Db 283 EQQHKSRQVDEAIGDLENLNDLEILPTNGEDEDGDDADIDD--EDIKSAIDNDEL 340
QY 180 FKERNGNHDPDMKAVIYSHKFWSGQDRSSADKRKYGDPAFPAGTGLVDMSDRN 239
Db 341 AKKYPVATSTTKVPTTLATSKTSSSSSSSTTTTMTATSTATSPSPSTTTTK----- 394
QY 240 IPRSPSPGEG-----FVNFDYGFAGTADADKTVWTHGNHVPAPN 282
Db 395 -PRTPT-PTIGRKLKNFTGLSPETRYGSDIGRIRAQTED-DTI-----BEKDK 444

QY 283 GSLGAMHVYESKFRNWSEGYSDFD-----RGAYVITFIPKSNWNTAPDKVKQ 328
Db 445 ENVGELHYDISGGSSSRKLVSFDPKSENYLSTYYPGKMN-ATEKKQK 493

RESULT 7
Q65311 PRELIMINARY; PRT; 736 AA.
ID O65311;
AC O65311;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE CAPSID PROTEIN.
OS adeno-associated virus 3.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OC NCBI_TaxID=46350;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3H;
RX MEDLINE=96266430; PubMed=8661429;
RA Muramatsu S., Mizukami H., Young N.S., Brown K.E.;
RT "Nucleotide sequencing and generation of an infectious clone of adeno-
RT associated virus 3";
RL Virology 221:208-217(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=3H;
RA Muramatsu S., Brown K.E.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL: U48704; AAC55049.1; .
DR InterPro: IPR001403; Parvo-coat.
DR Pfam: PF00740; Parvo-coat; 1.
SQ SEQUENCE 736 AA; 81660 MW; AFF1EF47B5C67A10 CRC64;

Query Match 6.1%; Score 110; DB 12; Length 736;
Best Local Similarity 22.4%; Pred. No. 4.9;
Matches 77; Conservative 49; Mismatches 144; Indels 74; Gaps 22;

QY 18 DPYRPSYGRAETVNNYIRKQVYSHRDGRKQMQTEQREWLSCYGV--GVTVNSGQ- 74
Db 4 DGYLPDW--LEDNLSEIGREW---WALKPGVPQKANKQHQDNRGLVPGYKYLGPNG 58
QY 75 ---YPTNRLAFASFDEDFRKNELKNG-RP--RSGETRAEFGRVAKESFDEKGFQRA 126
Db 59 LDKGEPVNEADAAAEHDKAYDQOLKAGDNPYLKYNHADADEFORLQE---DTSFGNGL 115
QY 127 REVASVMNRAL--NAHDESAYLDNLKKELANGNDALRNEDARSPPYSALRNTPSFKERN 184
Db 116 RAVQAKRILLEPLGLVEEAATAKTPGKK-----GAYDQSPQEPDSS 156
QY 185 GG---NHDPSPMKAVIYSHKFWSGQDRSSADKRKYGDPAFPAP-GTGLVDMSDRN 239
Db 157 SGVGKSGKQPKARKL-----NFGQTGDSVPDPQPLGEPPA---APTSLGSNTMASGG 208
QY 240 IPRSPSPG-----EGFVNFYGFAGTADADKTVW---THGNH-YHAPNGSLGAM 288
Db 209 APMAADNEGAGVNSNGNHCDSQLGDRVITTTT-WALPTYNNHLYKQISSQSGAS 267
QY 289 HVYESKFRNWSE--GYSDFRGAYVITFIPKSNWNTAPDKVKQGW 330
Db 268 N--DNHYFGYSTPWGYEDFNR--FHCHFSRDMQRL---INNNW 304

RESULT 8
Q07351 PRELIMINARY; PRT; 919 AA.
ID O07351;
AC O07351;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PROTEASE PRECURSOR.
GN PRTV.

OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=017;
RX MEDLINE=89013889; PubMed=3050359;
RA Alm R.A., Stroeder U.H., Manning P.A.;
RT "Extracellular proteins of Vibrio cholerae: Nucleotide sequence of the structural gene (hlyA) for the haemolysin of the haemolytic EL Tor strain 017 and characterization of the hlyA mutation in the non-haemolytic classical strain 569B";
RT Mol. Microbiol. 2:481-488(1986).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=017;
RA Manning P.A.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y00557; CAA68636.1; -
DR MEROPS; M06.002; -
DR InterPro; IPR000501; PKD_domain.
DR Pfam; PF00801; PKD; 2.
DR SMART; SM0089; PKD; 2.
DR PROSITE; PS0093; PKD; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW SIGNAL.
FT SIGNAL.
SQ SEQUENCE 919 AA; 101996 MW; 6A80774801FBD8D CRC64;

Query Match 6.0%; Score 109; DB 2; Length 919;
Best Local Similarity 21.8%; Pred. No. 7.8;
Matches 67; Conservative 52; Mismatches 111; Indels 78; Gaps 16;

QY 42 YSHR-DGRKQMTTEQREWLSCYGVTVNSGGQYPTNRL-AFASFDED-REFKNEKNGR 98
DB 446 YSNRGDGLNRSRP-----LTIPAGSQTARFKAWFOIEKXDYIARVLNGK 493

QY 99 PRSGETRAEF-----GRVAKESFDEKGFQAR-EVASVMNRALENADH-----ES 144
DB 494 PIAGNTTWDPPFKSLGSLVPAISSQSGGWVDAQFDLSAWAGQTVELAFDYLTDGGLAMEG 553

QY 145 AYLDNLKKELANGNDALRNEDARSFYSLRNTPSFK--ERNGNHDPSPRMKAVIYSHF 202
DB 554 LYVDDLRLEVDGNQTLIDNAEGTSSF-----AFQGTCKNGGFHE-----ANHY 596

QY 203 WSGODRSS-----ADKKRYGDPDAFRPAGTGLVDMSRDNIPRSPTSPGEGFVNFDY 256
DB 597 YLLOWSHNDVDOGLANLKRFFQLMSFEPLLVWYVDESADNWTGK--HPGE----- 647

QY 257 GWFGAGTEADADKTVTHGNHYPHAPNGSLGAMHYVESKFRNWSGYSDFDRGAYVITFIP 316
DB 648 GWLGG-RVDADQNALVMSKTG-----EVAQTRFQVRDATFSLFDQAP--LKLVT 692

QY 317 KSNWNTAPD 324
DB 693 AGDNTLED 700

RESULT 9
Q96C72 ID Q96C72 PRELIMINARY; PRT; 401 AA.
AC Q96C72;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DE HYPOTHETICAL 48.6 KDA PROTEIN (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE, AND RHABDOMYOSARCOMA;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014625; AAH14625.1; -
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 401 AA; 48615 MW; D9ADFDCC029A0851D CRC64;

Query Match 6.0%; Score 108; DB 4; Length 401;
Best Local Similarity 21.5%; Pred. No. 3.1;
Matches 68; Conservative 33; Mismatches 118; Indels 98; Gaps 12;

QY 39 QQVYSHRDGRKQMTTEQREWLSCYGVTVNSGGQYPTNRLAFASFDEDREFKNEKNGR 98
DB 7 EERNLLEERKRQKERR-----ITYREKEEEQRAEEQMLKERERERAERA 57

QY 99 PRSGETRAEFGRVAKESFDEKGFQAR-EVASVMNRALENADHESAYLDNLKKELANGN 158
DB 58 KREELR-EYQERVVKLEVERKKRQRELEIEERRR-----FEERLGD 102

QY 159 DALRNEDAR-----SPFYSALRNTPSFK--RNGGNHDPSPRMKAVIYSHFWS 204
DB 103 SSLSKDSRWGDRDSEGTWRKGPEDSEWRGPPKEKWRGEGRD----- 147

QY 205 GODRSSADK--RKYGDDP-----AFRP-----APGTGLVDMSRDNIPRSPTSPGEGFVN 253
DB 148 -EDSRHRDERPRRLGDDDEDEFSLRPDDDRVPRGMDDRGPRGPEDEDRSRRGADD 206

QY 254 FDYQWFGAQT-----ADADKTVTHGNHYPHAPNGSLGAMHYVESKFRNWSGYSDFDR 307
DB 207 DRPSWRNTDDRRPRLIADRGNWRHADDDRRPPRGL-----DEDR 248

QY 308 GAVYITFIPKSNWNTAPD 324
DB 249 G-----SWRTADE 256

RESULT 10
Q9H0G5 ID Q9H0G5 PRELIMINARY; PRT; 558 AA.
AC Q9H0G5;
DT 01-MAR-2001 (TRENBLrel. 16, Created)
DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE HYPOTHETICAL 66.4 KDA PROTEIN.
GN DKFPZ434K1421.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TESTIS;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weill B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Ausorge W., Boecker M., Bloeker H., Bauersachs S., Blum H.,
RA Lauber J., Buesterhoef A., Beyer A., Koehler K., Strack N.,
RA Mewes H.W., Ottenwaelder B., Obermaier B., Tampe J., Heubner D.,
RA Wambutt R., Korn B., Klein M., Poustka A.;
RT "Towards a Catalog of Human Genes and Proteins: Sequencing and Analysis of 500 Novel Complete Protein Coding Human cDNAs";
RT Genome Res. 11:422-435(2001).
RL EMBL; AL136806; CAB66740.1; -
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 558 AA; 66390 MW; 99B7BDBCDF06F98D CRC64;

Query Match 6.0%; Score 108; DB 4; Length 558;
Best Local Similarity 23.8%; Pred. No. 4.7;
Matches 59; Conservative 31; Mismatches 86; Indels 72; Gaps 11;

QY 36 RKQOVYSHRGRKQOQTEQREWLSCYGVVWNSQYPTNRLAFASDEDRFKNELK 95
DB 326 RDQENHYTDYRKDRSHRRE-----ASHRDSHWKREHQ 361
QY 96 NGRPSGTRAEFGRVAKESFDEEKGFRARE---VASVMNRALENAHDE---SAYLDN 149
DB 362 EDKPRADQR-ERSDVWKREKDRYSQRREQROQNDQNRPSRSEKKEEKSAAKEEH 420
QY 150 LK--KELANGDALRNEDAR-----SPFYSA-----LRNTPS 179
DB 421 MKVRKERYENDKTYDRKREKREVGVQSSERNQDRKRESSPNSAKKFLDQBSNKKRNMAK 480
QY 180 KVRGNHDPGRMKAVIYSKHFWS--GQDRSSADKKRYGDPDA---FRPATGTGLVDM 234
DB 481 DKERN--QEKPSNSSSLGAKHRLTEQGEKGEQER----PPEAVSKFAKRNNEETVMS 534
QY 235 SDRNIPR 242
DB 535 ADRYLAR 542
RESULT 11
O56139
ID O56139 PRELIMINARY; PRT; 736 AA.
AC O56139;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CAPSID PROTEIN VPL.
OS adeno-associated virus 3B.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=68742;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98080418; PubMed=9420229;
RA Rutledge E.A., Halbert C.L., Russell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus
RT (AAV) serotypes other than AAV type 2.";
RL J. Virol. 72:309-319(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Rutledge E.A., Russell D.W.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF028705; AAB95450.1; -;
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81906 MW; D523331AD5FD070F CRC64;

Query Match 5.9%; Score 107; DB 12; Length 736;
Best Local Similarity 22.5%; Pred. No. 8.1; Indels 62; Gaps 21;
Matches 76; Conservative 47; Mismatches 153;
QY 18 DYPFSGRAETVNNYIRKQOVYSHRGRKQOQTEQREWLSCYGVVWNSQY- 74
DB 4 DGYLPDW--LEDNLSEGIREW--NALKPGVPQPKANQOHDNRRLGLVPGYKYLPGCNG 58
QY 75 ----YPTNRLAFASDEDR-FKNEKNG-RP--RSGETRAEFGVAKESFDEEKGFORA 126
DB 59 LKQGEVNEADAALAHDKAYDQQLKAGDNPYLKYNHADAERLQEQ---DTSFGGMLG 115
QY 127 REVASVMNRALENAHDESAYLNLKELANGDALRNEDARSPPYSALRNTPSFKERN 186
DB 116 RAVFQAKKRILEPL-----GLVEEAKTAPGKRKRPVQDSP-----QEPDSSSGVGS 162
QY 187 NHDPSRMKAVIYSKHFWSQDRSSADKKRYGDPDAFRAP-GTGLVDMSRDRNTPRSPT 245
DB 163 GKQAPARKRL-----NFGQTGDSVPDPQLGEPPTA---APTSLGNTMASGGGAPMADN 214
QY 246 SPG-----EGFVNFYDYGWGAQTEADADKTVM---THGNH-YHAPNGSLGAMHYESK 294
DB 215 NEGADGVGNSGNHWCDSQWLGDRTVITSTRT-WALPTNNHLYKQISSQSGASN-DNH 271

QY 295 FRNWSE--GYSDFDRGAYVITFIPKSNWNTAPDKVKQGW 330
DB 272 YFGYSIPWGYFDFNR--FHCHFSRDPWQRL---INNWW 304
RESULT 12
O56137
ID O56137 PRELIMINARY; PRT; 736 AA.
AC O56137;
DT 01-JUN-1998 (TRENBLrel. 06, Created)
DT 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CAPSID PROTEIN VPL.
OS adeno-associated virus 6.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=68558;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98080418; PubMed=9420229;
RA Rutledge E.A., Halbert C.L., Russell D.W.;
RT "Infectious clones and vectors derived from adeno-associated virus
RT (AAV) serotypes other than AAV type 2.";
RL J. Virol. 72:309-319(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Rutledge E.A., Russell D.W.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF028704; AAB95450.1; -;
DR InterPro; IPR001403; Parvo_coat.
DR Pfam; PF00740; Parvo_coat; 1.
SQ SEQUENCE 736 AA; 81411 MW; 311217A089C565F5 CRC64;

Query Match 5.9%; Score 106.5; DB 12; Length 736;
Best Local Similarity 21.5%; Pred. No. 8.9; Indels 73; Gaps 21;
Matches 74; Conservative 50; Mismatches 147;
QY 18 DYPFSGRAETVNNYIRKQOVYSHRGRKQOQTEQREWLSCYGVVWNSQY- 75
DB 4 DGYLPDW--LEDNLSEGIREW--WDLKPGAPKPKANQOQDGRGLVPGYKYLPGFNG 58
QY 76 ----PTNRLAFASDEDR-FKNEKNG-RP--RSGETRAEFGVAKESFDEEKGFORA 126
DB 59 LDKGEFVNEADAALAHDKAYDQQLKAGDNPYLKYNHADAERLQEQ---DTSFGGMLG 115
QY 127 REVASVMNRALE--NAHDESAYLDNLKELANGDALRNEDARSPPYSALRNTPSFKERN 184
DB 116 RAVFQAKKRILEPLFGLVEEAKTAPGKR-----PVEQSPQEPDSS 156
QY 185 GG----NHDPSRMKAVIYSKHFWSQDRSSADKKRYGDPDAFRAPGTGLVDMSRDNI 240
DB 157 SGICKTQQQPAKKRL-----NFGQTGDSVPDPQLGEPPTA-TPA-AVGPTTMASGGGA 209
QY 241 PRSPTSPG-----EGFVNFYDYGWGAQTEADADKTVM---THGNHYH--APNGSLGAM 288
DB 210 PMADNNEGADGVGNASGNWHCHDSTWLGDRVITSTRT-WALPTNNHLYKQISSASTGAS 268
QY 289 HYESKFRNWE--GYSDFDRGAYVITFIPKSNWNTAPDKVKQGW 330
DB 269 N--DNHYFGYSTPWGYFDFNR--FHCHFSRDPWQRL---INNWW 305
RESULT 13
O9WB28
ID O9WB28 PRELIMINARY; PRT; 736 AA.
AC O9WB28;
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CAPSID PROTEIN.
OS adeno-associated virus 1.
OC Viruses; ssDNA viruses; Parvoviridae; Parvovirinae; Dependovirus.
OX NCBI_TaxID=85106;

```

RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=99214338; PubMed=10196295;
RA  Xiao W., Wilson J.M.;
RT  Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RL  EMBL; AF063497; AAD27757.1;
DR  InterPro; IPR001403; Parvo_coat.
DR  Pfam; PF00740; Parvo_coat; 1.
SQ  SEQUENCE 736 AA; 81375 MW; CFABFB9BD5CD0595 CRC64;

Query Match          5.9%; Score 106.5; DB 12; Length 735;
Best Local Similarity 21.5%; Pred. No. 8.9;
Matches 74; Conservative 50; Mismatches 147; Indels 73; Gaps 21;

QY  18 DVPKPSGRAETVYNNVIRKWOOVYSHRDGRKQCMTEEQREWLSYGCVGVTWVNSQY-- 75
DB  4 DGYLPDW--LEDNLSEIGIREW--WDLKPGAPKPKANQKQDDGRGLVLPGYKYLGFENG 58
QY  76 -----PTNRLAPASFDEDR-FKNELNG-RP--RSGETRAEFEGRVAKESFDEEKQFQA 126
DB  59 LDGEPVNAADAALAHDKAYDQOLKAGDNPYLRYNEADAEOERLQE---DTSFGNGLG 115
QY  127 REVASVNNRALE--NAHDESAYLNLKELANGNDALRNEDARSPPYSALRNTPPSKERN 184
DB  116 RAVFOAKKRVLEPLGLVEEGAKTAPGKKR-----PVEQSPQEPDSS 156
QY  185 GG-----NHDSPRKAVIYSHKFGSGQDRSSADKRYGDPDAFRPAPGTGLVDMSDRNI 240
DB  157 SGIGKGTQQPAKKRL-----NFGQTDGSEVPPQPLGEPPA-TPA-AVGPTTMSAGGGA 209
QY  241 PRSPTSPG-----EGFVNFDFGWFGAQTEADADKTW---THGNHYH--APNGLSLGAM 288
DB  210 PMADNNEGADGVGNASGNWCHDSTLWGLDRVITSTRT-WALPTYNHNLKQIASSATGAS 268
QY  289 HVYESKFRNWE--GYSDFRGAYVITFPKSWNTAPDKVKQGW 330
DB  269 N--DNHYFGYSTPWGYPDFNR--FCHFSRQWRL---INNWW 305

RESULT 14
Q9KM06 ID Q9KM06 PRELIMINARY; PRT; 918 AA.
AC Q9KM06;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PROTEASE.
GN VCA0223
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=566;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Emoliaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
DR EMBL; AE004362; AAF96135.1; -.
DR MROPS; M06.002; -.

```

```

DR TIGR; VCA0223; -.
DR InterPro; IPR000601; PKD_domain.
DR InterPro; IPR000130; ZP_Mtpeptide.
DR Pfam; PF00801; PKD; 2.
DR SMART; SM00089; PKD; 2.
DR PROSITE; PS00093; PKD; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 918 AA; 101882 MW; 4F874BAFBBD63FD07 CRC64;

Query Match          5.9%; Score 106.5; DB 16; Length 918;
Best Local Similarity 21.2%; Pred. No. 12;
Matches 65; Conservative 52; Mismatches 113; Indels 77; Gaps 16;

QY  42 YSHR-DGRKQKQMTREQLSYGCVGVTVWVNSGOVPTNRL-APASFDEE-REFKNELKNR 98
DB  446 YENRGDDLNKRNRP-----LTI FAGSQATFRFRANFQIEKDYARVVLNGK 493
QY  99 PRSGETRA---EFEGRVAKESFDEEKFORAR-EVASVNNRALENAHD-----ESA 145
DB  494 PIAGNTTMDDPFKGSLVPAISGGQSDGVDAQFDLSAWAGQTVLAFDYLTDGGLAMEGL 553
QY  146 YLDNLKELANGNDALRNEDARSPPYSALRNTPPFK--ERNGGNHDPSPRKAVIYSHKFW 203
DB  554 YVDDLRLEVDGNTLIDNAEGTSSP-----AFQGTKNGGPHE-----ANHYY 596
QY  204 SQDRSSS-----ADKRYGDPDAFRPAPGTGLVDMSDRNI PRSPTSPGEGFVNFYD 257
DB  597 LLQWRSHNDVDOGLANLRFQGLMSFEPGLLVWYVDESADNVWVGK--HPGE-----G 647
QY  258 WFGAQTADADKTWTHGNHYHAPNGSLGAMHVYESKFRNWSGYSDFDRGAYVITFPK 317
DB  648 WLGV-VDAQNALVWSKYG-----EVAQTRFQVRDAIFSLFDQAP--LKLVA 692
QY  318 SWNTAPD 324
DB  693 DGNLTED 699

RESULT 15
O88466 ID O88466 PRELIMINARY; PRT; 1888 AA.
AC O88466;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ZINC FINGER PROTEIN 106.
GN ZFP106 OR H3A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6;
RX Zuberi A.R., Christianson G.J., Mendoza L., Shastri N.,
RA Roopenian D.C.;
RT "The mouse H3a minor histocompatibility antigen, the cytotoxic
RT determinant of the H3 transplantation rejection locus, is encoded by a
RT novel zinc finger (C2H2) motif-containing gene, zfp106."
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 805-890 FROM N.A.
RX STRAIN=BALB/CJ AND C57BL/6 X DBA/2, 12/3, AND C57BL/6;
RX MEDLINE=99060924; PubMed=9846490;
RX Zuberi A.R., Christianson G.J., Mendoza L.M., Shastri N.,
RA Roopenian D.C.;
RT "Positional cloning and molecular characterization of an
RT immunodominant cytotoxic determinant of the mouse H3 minor
RT histocompatibility complex."
RL Immunity 9:687-698(1998).
CC -!- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).

```

```
DR EMBL; AF060246; AAD04329.1; -
DR EMBL; AF067399; AAD04342.1; -
DR EMBL; AF060245; AAD04328.1; -
DR EMBL; AF067398; AAD04341.1; -
DR EMBL; AF067397; AAD04340.1; -
DR MGD; MGI:104549; Zfp106.
DR InterPro; IPR001680; WD40.
DR InterPro; IPR000822; Znf-C2H2.
DR Pfam; PF00400; WD40. 6.
DR Pfam; PF00096; Zf-C2H2. 1.
DR PRINTS; PR00320; GPROTEINBRPT.
DR SMART; SM00320; WD40; 6.
DR SMART; SM00355; Znf-C2H2; 2.
DR PROSITE; PS50082; WD_REPEATS.2; 2.
DR PROSITE; PS50294; WD_REPEATS_REGION.1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2.1; 2.
KW DNA-binding; Metal-binding; Repeat; WD repeat; Zinc-finger.
FT CHAIN 556 664 H3A MINOR HISTOCOMPATIBILITY ANTIGEN.
SQ SEQUENCE 1888 AA; 208867 MW; 0216B92854698C58 CRC64;

Query Match          5.8%; Score 105.5; DB 11; Length 1888;
Best Local Similarity 19.6%; Pred. No. 37;
Matches 68; Conservative 37; Mismatches 101; Indels 141; Gaps 17;

QY      8 PPAELDRMPDPYRPSYGRAETVNNYIRKWOQVYSHRDGRKQOMTEQREWLSYGCVGV 67
DB      109 PPSNSQEVNSDDRQPCWRREDIPYQDRESYQPPRRHRG-----PPQDWM----- 154

QY      68 TWVNSQYPTNRLAFASFDEDEFRKNEKLN-GRPR-----SGETRA----- 106
DB      155 KWEKDG-----FNSTRKNSFPHSLRNSGGPGSSVYHKGATRGSGSTWFLNHSNSGG 206

QY      107 -----EFGRVAKESFDEE--KGFORAREVASYMNRALENAHDESAYLDNLKKEL 154
DB      207 WHSNNGMVDWNTNGTRNSWHSEGTGGF-----PSWEMNN----- 242

QY      155 ANGNDALRNEDARSPFYSAALNT-----PSFKERNGNHDPSSRM-----K 194
DB      243 SNGN-----WKS SVRSTNSWNYNGPGDKFQGGRRNPNYQMEDMTKMNKKS NK 291

QY      195 AVLYSKH--FWSGQDRSSSADKKYGDPAFPAPGTGLVDMSRDRNIPRSPTSPGEGFV 252
DB      292 PSKYSOERCKWQRDRDKAAKYR--SPPEGY-----ASDTPFSEGLL 331

QY      253 NFDYGFAGQTEADADKTVWTHGNHYHAPNGSLGAMHYVESKFRNWS 299
DB      332 EFNF----EQRESQTTKQDTPAASKINGKNGTKA-----ROKFRRW 369
```

Search completed: September 27, 2002, 12:40:32
Job time: 140 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 27, 2002, 12:38:12 ; Search time 13.4 Seconds
(without alignments)
956.431 Million cell updates/sec

Title: US-09-884-948-1
Perfect score: 1811
Sequence: 1 DSDRVTTPPAEPLDRMPDPY.....ITPIKSWNTAPDKVKQWGP 331

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1811	100.0	331	1 TGL_STRSS	P81453 streptococ
2	108	6.0	1382	1 IF3A_HUMAN	Q14152 homo sapien
3	104	5.7	575	1 AMVB_BACCI	P06547 bacillus ci
4	100.5	5.5	926	1 MAY3_SCHCO	P37934 schizosacch
5	99.5	5.3	939	1 ARS2_DROME	Q9V9K7 drosophila
6	99	5.5	2424	1 CCAA_RABIT	P27884 oryctolagus
7	98.5	5.4	1453	1 NKCR_MOUSE	P30415 mus musculu
8	98	5.4	350	1 PHOE_SALTY	P30705 salmonella
9	98	5.4	547	1 BX42_DROME	P39736 drosophila
10	98	5.4	1462	1 NKCR_HUMAN	P30414 homo sapien
11	96	5.3	350	1 PHOE_SALTY	Q56119 salmonella
12	96	5.3	536	1 SKIP_HUMAN	Q13573 homo sapien
13	96	5.3	1344	1 IF3A_MOUSE	P23116 mus musculu
14	96	5.3	2468	1 MAPB_HUMAN	P46821 homo sapien
15	95	5.2	572	1 LMD1_HUMAN	P29536 homo sapien
16	95	5.2	1253	1 MYSG_DROME	Q01989 drosophila
17	94	5.2	462	1 YFHD_HAETN	P44587 haemophilus
18	93.5	5.2	396	1 TRT_DROME	P19351 drosophila
19	93.5	5.2	1722	1 RB22_HUMAN	P29375 homo sapien
20	93	5.1	482	1 UR22_HUMAN	Q15696 homo sapien
21	92.5	5.1	633	1 MLH_TETTH	P40631 tetrahymena
22	92.5	5.1	790	1 KIF9_MOUSE	Q9WV04 mus musculu
23	92	5.1	1101	1 DIA2_HUMAN	Q60879 homo sapien
24	92	5.1	2459	1 MAPB_RAT	P15205 rattus norv
25	91.5	5.1	381	1 NCAP_CVCAI	P35298 canine ente
26	91.5	5.1	383	1 OMS2_SALTY	Q56111 salmonella
27	91	5.0	350	1 R340_ARATH	P92965 arabidopsis
28	91	5.0	535	1 YGH1_CAEEL	Q22836 caenorhabdi
29	91	5.0	636	1 GYRB_THEMEA	P77993 thermotoga
30	91	5.0	1536	1 SIN3_YEAST	P22579 saccharomyc
31	90.5	5.0	351	1 PHOE_KLEPN	P30704 klebsiella
32	90.5	5.0	367	1 OMPK_ECOLI	P06996 escherichia
33	90.5	5.0	980	1 BOB1_YEAST	P38041 saccharomyc

34 90.5 5.0 1131 1 YMS2_CAEEL P34498 caenorhabdi
35 90 5.0 351 1 PHOE_CITFR Q01605 citrobacter
36 90 5.0 611 1 IF4B_HUMAN P23588 homo sapien
37 89.5 4.9 537 1 ARP_PLAFA P04931 plasmodium
38 89.5 4.9 935 1 KINE_SINRA C43093 syncephalids
39 89.5 4.9 3421 1 TEGU_HSVB P28955 equine herp
40 89 4.9 977 1 DLPI_HUMAN O14490 homo sapien
41 88.5 4.9 532 1 CCBI_DROAC O44220 drosophila
42 88 4.9 544 1 DSKI_SCHPO P36616 schizosacch
43 88 4.9 770 1 DAB2_HUMAN P98082 homo sapien
44 87.5 4.8 264 1 RS31_ARATH P92964 arabidopsis
45 87.5 4.8 303 1 SFRI_ARATH O22315 arabidopsis

ALIGNMENTS

RESULT 1

TGL_STRSS
ID TGL_STRSS STANDARD; PRT; 331 AA.
AC P81453;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Protein-glutamine gamma-glutamyltransferase (EC 2.3.2.13)
DE (transglutaminase) (TGase).
DE Streptococcus sp. (strain S-8112).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxID=86037;
[1]
RN SEQUENCE.
RX MEDLINE=93280110; Pubmed=8099353;
RA Kanaji T., Ozaki H., Takac T., Kawajiri H., Ide H., Motoki M.,
RA Shimonishi Y.;
RT "Primary structure of microbial transglutaminase from
RT Streptococcus sp. strain S-8112.";
RL J. Biol. Chem. 268:11565-11572(1993).
CC -!- FUNCTION: CATALYZES THE CROSS-LINKING OF PROTEINS AND THE
CC CONJUGATION OF POLYAMINES TO PROTEINS.
CC -!- CATALYTIC ACTIVITY: Protein glutamine + alkylamine = protein N5-
CC alkylglutamine + NH(3).
KW Transferase; Acyltransferase.
FT ACT_SITE 64 64 POTENTIAL.
SQ SEQUENCE 331 AA; 37862 MW; 5992363A63B2C1FA CRC64;

Query Match 100.0%; Score 1811; DB 1; Length 331;
Best Local Similarity 100.0%; Pred. No. 4.5e-129;
Matches 331; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 DSDRVTTPPAEPLDRMPDPYRPSYGRAETVYNNYIRKQQVYSHRDGRKQOMTEEQREW 60
Db 1 DSDRVTTPPAEPLDRMPDPYRPSYGRAETVYNNYIRKQQVYSHRDGRKQOMTEEQREW 60
Qy 61 SYGCVGVTWNSGGYPTNRLAFASDFEDRKNELKNGRPSGGETRAFEGRVAKESFDEE 120
Db 61 SYGCVGVTWNSGGYPTNRLAFASDFEDRKNELKNGRPSGGETRAFEGRVAKESFDEE 120
Qy 121 KGFORAREVASVMNRALENAHDESAYLDNLKELANGNDALRNEDARSFYSLALRNTPSF 180
Db 121 KGFORAREVASVMNRALENAHDESAYLDNLKELANGNDALRNEDARSFYSLALRNTPSF 180
Qy 181 KERNGGHDPSPRMKAVLYSKHFWSGQDRSSADKKYGDPAFPAPGCTGLVDMSDRNI 240
Db 181 KERNGGHDPSPRMKAVLYSKHFWSGQDRSSADKKYGDPAFPAPGCTGLVDMSDRNI 240
Qy 241 PRSPTSPGEGFVNFDYGNWFGAQTEADADKVTWTHGNHYHAPNGSLGAMHYESKFRNWE 300
Db 241 PRSPTSPGEGFVNFDYGNWFGAQTEADADKVTWTHGNHYHAPNGSLGAMHYESKFRNWE 300
Qy 301 GYSDFDGRGAVYITPIPKSWNTAPDKVKQWGP 331
|||||

```

Db 301 GYSDRGAYVITFPKSWNTAPDKVKGWP 331
RESULT 2
IF3A_HUMAN
ID IF3A_HUMAN STANDARD; PRT: 1382 AA.
AC Q14152; O00653;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE Eukaryotic translation initiation factor 3 subunit 10 (eIF-3 theta)
DE (eIF3 p167) (eIF3 p180) (eIF3 p185).
GN EIF3S10 OR KIAA0139.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RX MEDLINE=96127530; PubMed=8590280;
RA Nagase T., Seki N., Tanaka A., Ishikawa K.-I., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. IV.
RT The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
RT analysis of cDNA clones from human cell line KG-1.";
RL DNA Res. 2:167-174(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97294683; PubMed=9150439;
RA Scholler J.K., Kanner S.B.;
RT "The human p167 gene encodes a unique structural protein that contains
RT translation initiation factor 3. Comparison of homologues from human,
RT Nicotiana tabacum, Caenorhabditis elegans, and Saccharomyces
RT cerevisiae.";
RL J. Biol. Chem. 272:7106-7113(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97207269; PubMed=9054404;
RA Johnson K.R., Merrick W.C., Zoll W.L., Zhu Y.;
RT "Identification of cDNA clones for the large subunit of eukaryotic
RT translation initiation factor 3. Comparison of homologues from human,
RT Nicotiana tabacum, Caenorhabditis elegans, and Saccharomyces
RT cerevisiae.";
RL J. Biol. Chem. 272:7106-7113(1997).
CC -!- FUNCTION: BINDS TO THE 40S RIBOSOME AND PROMOTES THE BINDING OF
CC METHIONYL-TRNAI AND MRNA.
CC -!- SUBUNIT: EIF-3 IS COMPOSED OF AT LEAST 10 DIFFERENT SUBUNITS.
CC -!- PTM: PHOSPHORYLATED.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: BELONGS TO THE EIF3S10 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D50923; BAA09488.1; -
DR EMBL; U58046; ABA41584.1; -
DR EMBL; U78311; AAB80895.1; -
DR MIM; 602039; -
DR InterPro; IPR000717; PCI.
DR InterPro; IPR002017; Spectrin.
DR SMART; SM00088; PINT; 1.
KW Initiation factor; Protein biosynthesis; Repeat; Phosphorylation.
FT DOMAIN 925 1172
FT [CP]-[PS]-[RW]-R-[GN]-[AM]
FT SEQUENCE 1382 AA; 166568 MW; 485C01828D67EBBA CRC64;
Query Match 6.0%; Score 108; DB 1; Length 1382;
Best Local Similarity 21.5%; Pred. No. 2.2;
Matches 68; Conservative 33; Mismatches 118; Indels 98; Gaps 12;

```

```

QY 39 QQVYSHRDGRKQOQMTTEQREWLISYGCQVTVWNSGOYPTNRLAFAFSDDEDRKFKNELKNGR 98
Db 785 EERNRLLEKRRQKERR-----ITYYEKEEERAEQMLKEREERAEERA 835
QY 99 PRSGETRAEFEGRAVESDEDEKGFQFARREVASVMRALENHADESAYLDNKKELAN 158
Db 836 KREEELR-EYGERVKKLEEVERKKRQRELEIERERR-----EEERLGD 880
QY 159 DALRNEDAR-----SPFYSLRNTPSFKE-RNGGNHDPSPMKAVIYSKFW 204
Db 881 SSLSRDSRWGDRDSEGTWRKGEADSEWRGPPPEKWEWRGEGRD----- 925
QY 205 GQDRSSADK---RKYGDDP-----APGGLVDMRDNRINPSPTSPGEGFVN 253
Db 926 -EDRSHRRDERPRRLGDDDEDEPSLRPDQDRVPRGMDDDRGRGPRGPRGPRGADD 984
QY 254 FDYGVFGAQT-----ADADKTWTHGNHYHAPNGSLGAMHYVESKFRNWSGEGYDFDR 307
Db 985 DRPSWRNTDDRPRIADDEGWRHADDPRRGL-----DEDR 1026
QY 308 GAVVITFPKSWNTAPD 324
Db 1027 G-----SWRTADE 1034
RESULT 3
AMVB_BACCI
ID AMVB_BACCI STANDARD; PRT: 575 AA.
AC P06547;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Beta-amylase precursor (EC 3.2.1.2) (1,4-alpha-D-glucan
DE maltohydrolase).
OS Bacillus circulans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCIB 11033;
RX MEDLINE=88260890; PubMed=2455212;
RA Siggens K.W.;
RT "Molecular cloning and characterization of the beta-amylase gene from
RT Bacillus circulans.";
RL Mol. Microbiol. 1:86-91(1987).
CC -!- FUNCTION: CATALYZES THE LIBERATION OF MALTOSE FROM 1,4-ALPHA-D-
CC GLUCANS.
CC -!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-alpha-glucosidic linkages in
CC polysaccharides so as to remove successive maltose units from the
CC non-reducing ends of the chains.
CC -!- SUBUNIT: MONOMER.
CC -!- SIMILARITY: BELONGS TO FAMILY 14 OF GLYCOSYL HYDROLASES
CC (BETA-AMYLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y00523; CAA68578.1; -
DR PIR; S03745; S03745.
DR HSP; P36924; I892.
DR InterPro; IPR001554; Glyco_hydro_14.
DR Pfam; PF01373; Glyco_hydro_14; 1.
DR PRINTS; PR00750; BETAAMYLASE.
DR PROSITE; PS00506; BETA_AMYLASE_1; 1.
DR PROSITE; PS00679; BETA_AMYLASE_2; 1.
KW Hydrolase; Glycosidase; Polysaccharide degradation; Signal.

```

```

FT SIGNAL 1 36
FT CHAIN 37 575 BETA-AMYLASE.
FT ACT_SITE 125 BY SIMILARITY.
FT ACT_SITE 199 BY SIMILARITY.
SQ SEQUENCE 575 AA: 62899 MW: 724B3C0D6B4A258 CRC64;

Query Match 5.7% Score 104; DB 1; Length 575;
Best Local Similarity 22.9% Pred. No. 1.5;
Matches 76; Conservative

QY 24 YGRAETVNNYRKQOVQYSHRDGRKQOMTEEQREWLSYGCVGVTW-----NSG 73
DB 80 WGYVESAGDNQF-DWSYKKTIVADYVKQ-----AGLKWPIISTRHCGGNVG 124
QY 74 Q-----YPTNRLAFASDEDFKNEKNGRPRSG-----ETRAEPEGVAKESFDEEXGFQ 124
DB 125 DDNCNIPLPWSLWKSADAMQFKE-----SGYVNNEISLPWSGVGKQ-YDE----- 171
QY 125 RAREVASVMNRALENADHESAYLONLAKKELANGNDALRNEDARSFYALSALNTPSFKERN 184
DB 172 ---LYASF-----AQNFSAYKDMIPKXILSGGP-----SGELRYPSTYPAA 209
QY 185 GGNHDPGRMKAVIYSKHFWSQDRSSADKKRYGDDPAFRPAGPTGLVDM3DRNRINPRSP 244
DB 210 GWSY-PARGKQVYTE---TAKSAFRAMTKYGLSKDKINAAGCNTLTSMSQ-----ISP 260
QY 245 TSPEGP-----VNFEDVG-----WF-----GAQTEADADKTVWTH-----GNH 277
DB 261 PTSDGGYTGGGYNTITYKDFLSYQSVLENHGLGVGAAAHKNFDPVGVIRIGAKISGIH 320
QY 278 YHAFNGSL--GAMHVYESKERNWSEGSDFDR 307
DB 321 WOMNPNMPHSAEH-----AGGYDYNR 343

RESULT 4
MAY3_SCHCO STANDARD; PRT; 926 AA.
ID MAY3_SCHCO
AC P37934;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mating-type protein A-alphaY3.
OS Schizopyllium commune (Bracket fungus).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Stereales; Schizophyllaceae; Schizophyllium.
OC NCBI_TaxId=5334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UVM 9-4;
RX MEDLINE=92357793; Pubmed=1353886;
RX Stankis M.M., Specht C.A., Yang H., Giasson L., Ullrich R.C.,
RA Novotny C.P.;
RT "The A alpha mating locus of Schizopyllium commune encodes two
RL dissimilar multiallelic homeodomain proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:7169-7173(1992).
CC -!- FUNCTION: SPECIFIES A-ALPHA-3 MATING-TYPE. MAY REGULATE THE
CC EXPRESSION OF GENES SPECIFIC TO THE HOMOKARYOTIC CELL TYPE.
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED CONSTITUTIVELY IN HOMOKARYONS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M97180; AAB01370.1;
CC PIR; B37271; B37271.
CC HSSP; P02833; 9ANT.

```

Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W., Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K., Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C., Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarly C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Palazzolo M., Pittman G.S., Pan S., Pollard J.J., Puri V., Reese M.G., Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong E., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstein G.N., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
The genome sequence of *Drosophila melanogaster*;
Science 287:2185-2195(2000).
-!- SIMILARITY: BELONGS TO THE ARS2 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF003784; AAF57281.1;
CC FlyBase; FBgn0033062; CG7843.
CC Hypothetical protein.
CC
CC SEQUENCE 939 AA; 106804 MW; 2AD3D528619AC08B CRC64;
Query Match 5.5%; Score 99.5; DB 1; Length 939;
Best Local Similarity 18.3%; Pred. No. 5.9;
Matches 69; Conservative 43; Mismatches 144; Indels 121; Gaps 13;
QY 10 AEPLDRPDPYRSYGAETVWNNYIRKQOVYSHRDGRKQMTBQREWLVSVCVGVTV 69
DB 311 ADPVSTQKVPVP-----VNSDGENWDDDAENSAFKKELAEKSD----- 351
QY 70 VNSGQYPTNKLAFASDEDFKLNELNGRPR-----SGTRAEFGRVAKESFDEE 120
DB 352 --SDSKP-----EDQLNKKTKKRNSSDDSSSSSSSSSSDEEKLKRYDYE 399
QY 121 KGFQREVASVKNRALENADHESAYLDNLKELANGNDALRNEDARSPYSALRNTPSF 180
DB 400 DGLRAEQKTEAEKDRQATKAKGQPSFKLDEDEGNEPTPKGLDSKINTYEIDNLTLS 459
QY 181 KE-----RNGGNHDPGRM-----KAVIYSKHFWSGODR-----SSADKRRKYGDPAFR 224
DB 460 PEISSNPIKNTDNGDSKVEDEGKPSV-----GDKVKVETETIDLDKVDGQPRALH 512
QY 225 PAPGTGLVMSRDNIPRST-SPGEGFVN-----FDYGFWFGAQTE 264
DB 513 RTSIFL-----RNLAPSITRTEIAVNCNRFSGYLVAITADPLVRRVTRGWIITFMRD 566
QY 265 ADADKTVYTHGN-----HYHAPNGSLGAMHYES----- 293
DB 567 VNKEICWGLNNQRLDCMGAIVNRDLRRVRPANGITAHKQVRSRDIKCAKIALNLD 626
QY 294 -KFRWSEGSDFDRGA 309
DB 627 EKFLWAEGPKDDNSA 643

RESULT 6
ID CCAA_RABIT STANDARD; PRT; 2424 AA.
AC P27884; P27883;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Voltage-dependent P/Q-type calcium channel alpha-1A subunit (Calcium
DE channel, L type, alpha-1 polypeptide isoform 4) (Brain calcium channel
DE 1) (Bi).
GN CACNA1A OR CACNL1A4 OR CACH4 OR CACN3.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RC MEDLINE=91187110; PubMed=1849233;
RA Mori Y., Friedrich T., Kim M.-S., Mikami A., Nakai J., Ruth P.,
RA Bosse E., Hofmann F., Flockerzi V., Furuichi T., Mikoshiba K.,
RA Imoto K., Tanabe T., Numa S.;
RA "Primary structure and functional expression from complementary DNA
RA of a brain calcium channel";
RA Nature 350:398-402(1991).
RN [2]
RP BETA-SUBUNIT BINDING DOMAIN, AND MUTAGENESIS.
RX MEDLINE=94150724; PubMed=7509046;
RA Pragnell M., de Waard M., Mori Y., Tanabe T., Snutch T.P.,
RA Campbell K.P.;
RA "Calcium channel beta-subunit binds to a conserved motif in the I-II
RA cytoplasmic linker of the alpha 1-subunit";
RA Nature 368:67-70(1994)
CC -!- FUNCTION: VOLTAGE-SENSITIVE CALCIUM CHANNELS (VSCC) MEDIATE THE
CC ENTRY OF CALCIUM IONS INTO EXCITABLE CELLS AND ARE ALSO INVOLVED
CC IN A VARIETY OF CALCIUM-DEPENDENT PROCESSES, INCLUDING MUSCLE
CC CONTRACTION, HORMONE OR NEUROTRANSMITTER RELEASE, GENE EXPRESSION,
CC CELL MOTILITY, CELL DIVISION AND CELL DEATH. THE ISOFORM ALPHA-1A
CC GIVES RISE TO P AND/OR Q-TYPE CALCIUM CURRENTS. P/Q-TYPE CALCIUM
CC CHANNELS BELONG TO THE "HIGH-VOLTAGE ACTIVATED" (HVA) GROUP AND
CC ARE BLOCKED BY THE FUNNEL TOXIN (FTX) AND BY THE OMEGA-AGATOXIN-
CC IVA (OMEGA-AGA-IVA). THEY ARE HOWEVER INSENSITIVE TO
CC DIHYDROPIRIDINES (DHP), AND OMEGA-CONOTOXIN-GVIA (OMEGA-CTX-
CC GVIA).
CC -!- SUBUNIT: VOLTAGE-DEPENDENT CALCIUM CHANNELS ARE MULTISUBUNIT
CC COMPLEXES, CONSISTING OF ALPHA-1, ALPHA-2, BETA AND DELTA SUBUNITS
CC IN A 1:1:1:1 RATIO. THE CHANNEL ACTIVITY IS DIRECTED BY THE PORE-
CC FORMING AND VOLTAGE-SENSITIVE ALPHA-1 SUBUNIT. IN MANY CASES, THIS
CC SUBUNIT IS SUFFICIENT TO GENERATE VOLTAGE-SENSITIVE CALCIUM
CC CHANNEL ACTIVITY. THE AUXILIARY SUBUNITS BETA AND ALPHA-2/DELTA
CC LINKED BY A DISULFIDE BRIDGE REGULATE THE CHANNEL ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- ALTERNATIVE PRODUCTS: IN THE BRAIN, A SHORT ISOFORM BI-1/1A-1 AND
CC A LONG ISOFORM BI-2/1A-2 (SHOWN HERE), ARE PRODUCED BY ALTERNATIVE
CC SPLICING
CC -!- TISSUE SPECIFICITY: BRAIN-SPECIFIC. PURKINJE CELLS CONTAIN
CC PREDOMINANTLY P-TYPE VSCC, THE Q-TYPE BEING A PROMINENT CALCIUM
CC CURRENT IN CEREBELLAR GRANULE CELLS.
CC -!- DOMAIN: EACH OF THE FOUR INTERNAL REPEATS CONTAINS FIVE
CC HYDROPHOBIC TRANSMEMBRANE SEGMENTS (S1, S2, S3, S5, S6) AND ONE
CC POSITIVELY CHARGED TRANSMEMBRANE SEGMENT (S4). S4 SEGMENTS
CC PROBABLY REPRESENT THE VOLTAGE-SENSOR AND ARE CHARACTERIZED BY A
CC SERIES OF POSITIVELY CHARGED AMINO ACIDS AT EVERY THIRD POSITION.
CC -!- SIMILARITY: BELONGS TO THE CALCIUM CHANNEL ALPHA-1 SUBUNITS
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

```

-----
EMBL: X57477; CAA40715.1; -
DR ENBL: X57689; CAA40872.1; -
DR ENBL: X57478; CAA40714.1; -
DR ENBL: X57688; CAA40871.1; -
DR InterPro: IPR000637; AT_hook.
DR InterPro: IPR002077; Ca_channel.
DR InterPro: IPR000636; Cat_chan_non_lig.
DR InterPro: IPR001682; Channel_pore_Ca_Na.
DR Pfam: PF00520; ion_trans_4.
DR PRINTS: PRO0167; CACHANNEL.
DR SMART: SM00384; AT_hook; 1.
KW Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KW Calcium channel; Glycoprotein; Repeat; Multigene family;
KW Calcium-binding; Phosphorylation; Alternative splicing.
FT REPEAT 83 363 I.
FT REPEAT 473 717 II.
FT REPEAT 1240 1523 III.
FT REPEAT 1560 1823 IV.
FT DOMAIN 1 98 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 99 117 S1 OF REPEAT I (POTENTIAL).
FT DOMAIN 118 135 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 136 155 S2 OF REPEAT I (POTENTIAL).
FT DOMAIN 156 167 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 168 185 S3 OF REPEAT I (POTENTIAL).
FT DOMAIN 186 190 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 191 209 S4 OF REPEAT I (POTENTIAL).
FT DOMAIN 210 228 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 229 248 S5 OF REPEAT I (POTENTIAL).
FT DOMAIN 249 335 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 336 360 S6 OF REPEAT I (POTENTIAL).
FT DOMAIN 361 487 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 488 506 S1 OF REPEAT II (POTENTIAL).
FT DOMAIN 507 521 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 522 541 S2 OF REPEAT II (POTENTIAL).
FT DOMAIN 542 549 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 550 568 S3 OF REPEAT II (POTENTIAL).
FT DOMAIN 569 578 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 579 597 S4 OF REPEAT II (POTENTIAL).
FT DOMAIN 598 616 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 617 636 S5 OF REPEAT II (POTENTIAL).
FT DOMAIN 637 689 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 690 714 S6 OF REPEAT II (POTENTIAL).
FT DOMAIN 715 1253 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1254 1272 S1 OF REPEAT III (POTENTIAL).
FT DOMAIN 1273 1288 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1289 1308 S2 OF REPEAT III (POTENTIAL).
FT DOMAIN 1309 1320 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1321 1339 S3 OF REPEAT III (POTENTIAL).
FT DOMAIN 1340 1350 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1351 1369 S4 OF REPEAT III (POTENTIAL).
FT DOMAIN 1370 1388 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1389 1408 S5 OF REPEAT III (POTENTIAL).
FT DOMAIN 1409 1495 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1496 1520 S6 OF REPEAT III (POTENTIAL).
FT DOMAIN 1521 1575 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1576 1604 S1 OF REPEAT IV (POTENTIAL).
FT DOMAIN 1605 1609 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1610 1629 S2 OF REPEAT IV (POTENTIAL).
FT DOMAIN 1630 1637 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1638 1656 S3 OF REPEAT IV (POTENTIAL).
FT DOMAIN 1657 1665 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1666 1684 S4 OF REPEAT IV (POTENTIAL).
FT DOMAIN 1685 1703 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1704 1723 S5 OF REPEAT IV (POTENTIAL).
FT DOMAIN 1724 1795 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1796 1820 S6 OF REPEAT IV (POTENTIAL).
FT DOMAIN 1821 2424 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 13 18 POLY-GLY.
FT TRANSMEM 727 732 POLY-GLU.
FT DOMAIN 1004 1010 POLY-GLY.
FT TRANSMEM 1012 1017 POLY-ARG.

```

```

FT DOMAIN 2219 2227 POLY-HIS.
FT DOMAIN 2242 2246 POLY-ARG.
FT DOMAIN 2288 2297 POLY-ARG.
FT DOMAIN 2298 2301 POLY-GLY.
FT DOMAIN 2372 2377 POLY-PRO.
FT DOMAIN 2411 2416 POLY-GLY.
FT TRANSMEM 383 400 BINDING TO THE BETA SUBUNIT.
FT SITE 318 318 CALCIUM ION SELECTIVITY AND PERMEABILITY
FT SITE 568 568 (BY SIMILARITY).
FT SITE 1469 1469 CALCIUM ION SELECTIVITY AND PERMEABILITY
FT SITE 1765 1765 (BY SIMILARITY).
FT MOD_RES 1831 1831 PHOSPHORYLATION (BY CAPK) (POTENTIAL).
FT CA_BIND 1849 1860 BY SIMILARITY.
FT CARBOHYD 1665 1665 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 772 1051 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 772 1120 MISSING (IN ISOFORM CBP103).
FT VARSPLIC 1857 1884 LYRDMVAMLRHMPPLGLGKNCARVAY -> HYKDMYSL
FT VARSPLIC 2230 2273 RVIISPLGLGKKCPHRVAC (IN ISOFORM
FT VARSPLIC 2273 CBP101/CBP109).
FT VARSPLIC 2273 RGPVRSVGSVARRRRRSGVAVRVPAPAPALAHARARAP
FT VARSPLIC 2273 ARK -> PAADAKRIQPDPRDGHGRARDQWRSRPS
FT VARSPLIC 2273 EGREHTTHRQ (IN ISOFORM BI-1/IA-1).
FT VARSPLIC 2273 MISSING (IN ISOFORM BI-1/IA-1).
FT VARSPLIC 2273 A -> T (IN ISOFORM CBS).
FT VARSPLIC 2273 S -> N (IN ISOFORM CBS).
FT VARSPLIC 2273 E -> S; REDUCED BETA-SUBUNIT INTERACTION.
FT VARSPLIC 2273 L -> H; REDUCED BETA-SUBUNIT INTERACTION.
FT VARSPLIC 2273 Y -> S; REDUCED BETA-SUBUNIT INTERACTION.
FT VARSPLIC 2273 E -> A; NO EFFECT ON BETA-SUBUNIT
FT VARSPLIC 2273 INTERACTION.
FT VARSPLIC 2273 SEQUENCE 2424 AA; 273228 MW; F7CC4D0AB4B45604 CRC64;

Query Match 5.5%; Score 99; DB 1; Length 2424;
Best Local Similarity 26.5%; Pred. No. 21;
Matches 43; Conservative 18; Mismatches 61; Indels 40; Gaps 7;

QY 99 PRSGTAEFEGRVAKESFDEKGFQRAREVASVMNRALENAHDESAYLDNLKKELANGN 158
Db 828 PQENNNNTNKSVAEPTVDQLGQRAEDFLRQARHHRARDPSAH----- 875
QY 159 DALRNEDARSPYSA-----LRNTSFKE-----RNGNHDPSRKAVIYKHFWSGQDR 208
Db 876 -AAAGLDARRPWAGSQEALSREGPYGREGSDHQAREGLEPPG-----FWEGE-- 922
QY 209 SSSADKRKYDPAFRPAPGTGLVDMSRDN-IPRSPTSPGE 249
Db 923 ---AEEGRAGDPHR-RHAHQGVGGSGSGSRGSPRTGTADGE 960

RESULT 7
NCRC_MOUSE
ID NCRC_MOUSE STANDARD; PRT; 1453 AA.
AC P30415;
DT 01-APR-1993 (Rel. 25, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE NK-tumor recognition protein (Natural-killer cells cyclophilin-
GN NKTR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93133824; PubMed=8421688;

```

RA Anderson S.K., Gallinger S., Roder J., Frey J., Young H.A.,
RT Ortaldo J.R.;
RT "A cyclophilin-related protein involved in the function of natural
RT killer cells."; Proc. Natl. Acad. Sci. U.S.A. 90:542-546(1993).
RN [2]
RN REVISIONS TO C-TERMINUS.
RC STRAIN=BALB/C; TISSUE=Blood;
RA Anderson S.K.;
RA Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: COMPONENT OF A PUTATIVE TUMOR-RECOGNITION COMPLEX.
CC INVOLVED IN THE FUNCTION OF NK CELLS.
CC -1- SIMILARITY: CONTAINS 1 CYCLOPHILIN-LIKE PPIASE DOMAIN.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L04289; AAA37500.2; ALT_INIT.
CC HSSP; Q27450; 1A33.
CC MGD; MGI:97346; Nktr.
CC InterPro: IPR002130; CSA_PPIase.
CC Pfam: PF00150; pro_isomerase; 1.
CC PRINTS; PR00153; CSA_PPIase.
CC PROSITE; PS00707; CSA_PPIase_1; 1.
CC PROSITE; PS00702; CSA_PPIase_2; 1.
CC Cyclosporin; Isomerase; Rotamase; Repeat; Transmembrane.
FT DOMAIN 1 176 PPIASE, CYCLOPHILIN-TYPE.
FT DOMAIN 222 241 ARG/LYS-RICH (BASIC).
FT DOMAIN 422 459 ARG/LYS-RICH (BASIC).
FT DOMAIN 964 1003 ARG/LYS-RICH (BASIC).
FT DOMAIN 198 273 ARG/SER-RICH.
FT DOMAIN 468 565 ARG/SER-RICH.
FT DOMAIN 658 842 ARG/SER-RICH.
FT DOMAIN 1303 1453 ARG-SER TANDEM REPEAT-RICH.
FT SEQUENCE 1453 AA; 163439 MW; DF1173FF814B283E CRC64;
-SQ

Query Match 5.4%; Score 98.5; DB 1; Length 1453;
Best Local Similarity 17.7%; Pred No. 12;
Matches 88; Conservative 63; Mismatches 152; Indels 195; Gaps 21;
QY 1 DSDDRVTPP--AEPLDMPDPYRPSYGRA-----ETVYNNYIRKQQ-- 40
Db 355 DDDSETPPHKEEMQL-RAYRPPSGEKWSKGDKLSDFCSSRWDRSLQSRSRWSYNG 413
QY 41 VYS-----HRDG-----RKQMTTEOREWLSYCGVGTWYN 71
Db 414 YSDLTARESDGHKKHKKRKEKFKKAKKOKHCHRRHQTAKRR-----IVMPDLE 466
QY 72 SGQYPTNRL-----AFASFDDEDRFXNELNCRPSRGETRAEFEGVRAKE 115
Db 467 PRRSPTRHMKSCVRRERRSRASSSSHHSSKRWKSDQDDGSASTHSSRDSV--RSKSH 524
QY 116 SPFEKGFQFARAVSYVMNPALENADESA-----YLDNLKELANGDALNEDAR 167
Db 525 SRSDSGSSRSRAVSKSSSSSLNRKSRSSRSRSGPRRTSISPKKPAQLSENKPVKTEPLR 584
QY 168 -----SPFYSALR----- 175
Db 585 PSVPQNGNLVQVAAENIPVILSDSPSPRWKPKQKPKPSYRIQEMKAKTTHLLPV 644
QY 176 -----NTPFKERNGNHDPSRMKAVISKFWSCQDPSRSSADKRYCDPDA 222
Db 645 QSTYSLTNIKATVSSSSYHKR-----BKPSESDGSAYSKY-----SDRSSGSSGR--SGSKS 694
QY 223 FRPAPGTGLVDMRDRNP-----RSPT-----SPGEGFVNFYGFQATADAD-- 268
Db 695 SRGSSRSRSTRSRSLPSTRSLSRSPSRSHSPKNSYDGSQSHSRSSSVTSVSSDDGR 754

QY 269 -----KTWTEGHNHAPNGSLGAMHV-----YESKERNWSEGYSDPD-----RGAYV 311
Db 755 AMFRSNRKKSVTSKRRHSNSEKTLHSKYVGRKSSRHRKYSERSSLDYTSDDSHV 814
QY 312 ITFIPKSWNTAPDKVKQG 329
Db 815 QVY-----SAPEKEKQG 826
RESULT 8
PHOE_SALTY STANDARD; PRT; 350 AA.
ID PHOE_SALTY
AC P30705;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Outer membrane pore protein E precursor.
GN PHOE OR STM0320.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
FN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RX MEDLINE=93083994; PubMed=1280609;
RA Sierings G., Elders R., van Lith B., Hofstra H., Tommassen J.;
RT "Characterization of the Salmonella typhimurium phoE gene and
RT development of Salmonella-specific DNA probes.";
RL Gene 122:45-52(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -1- FUNCTION: THIS IS ONE OF THE PROTEINS INDUCED WHEN CELLS ARE
CC GROWN UNDER PHOSPHATE LIMITATION. ITS PROTEIN PORE IS
CC PARTICULARLY EFFICIENT IN THE UPTAKE OF INORGANIC PHOSPHATE,
CC PHOSPHORYLATED COMPOUNDS, AND SOME OTHER NEGATIVELY CHARGED
CC SOLUTES.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC -1- SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X68023; CAA48164.1; -
CC EMBL; AE008709; AAL19276.1; -
CC PIR; S25525; S25525.
CC HSSP; P02932; 1PHO.
CC StyGene; SG10291; phoE.
CC InterPro; IPR001702; Gram_neg_porin.
CC Pfam; PF00267; Gram_ve_porins; 1.
CC PRINTS; PR00182; ECOLNEIPORIN.
CC PROSITE; PS00576; GRAM_NEG_PORIN; 1.
KW Outer membrane; Transmembrane; Porin; Signal; Complete proteome.
FT SIGNAL 1 20
FT CHAIN 21 350 OUTER MEMBRANE PORE PROTEIN E.
FT SEQUENCE 350 AA; 38762 MW; 9D1EE3355AF59877 CRC64;
SQ

```

Query Match          5.4%; Score 98; DB 1; Length 350;
Best Local Similarity 20.8%; Pred. No. 2.2;
Matches 64; Conservative 35; Mismatches 116; Indels 92; Gaps 13;

QY 24 YGRAETVYNNYIRKQVQYSHRDGRKQMTTEQREWLSCVGVTVWNS-----GOYPTN 78
DB 34 YGKVKAM-----HYMSYDSKDG-----DSYVRFKGETQINDQLTGVRWEAE 79
QY 79 RIASFDEFDRFKNELKNGRPSGTRAEFEGRVAKA--SFDEKGFQREAVASVMNRA 136
DB 80 ---FASNAESDSSOOK-----TRLAFAGLKLKDIGSPDYGRNLGALYDVEAWTDMF 128
QY 137 LENAHDESAYLDNLKELANGNDALRNEDARSPFYSALRNTPSPFKERNGNHDPSPRMKAV 196
DB 129 PFGGDSRAQTDNFWTKKASGLAYRNTD-----PFGIVDGLDLILQYCKNEDRD----- 179
QY 197 YSKFHWGSGQDRSSADRRKRYGDDPAFPAGTGLVDMRDNRNIPRSTSPGEGFVNF 256
DB 180 -VKKONGDGFSTSYD-----FEGSD-FAVSGAYTILSDRTREQNLQRGT----- 223
QY 257 GWFGAQTADADKTWTWTHGNHYHAPNGSLGAMHYVESKERNWSEGSYDFDYGAVYIFIP 316
DB 224 ---GDKAE-----WATGVKYDA-----NDIITAFYS 248
QY 317 KSWNTAP 323
DB 249 ETRNMT 255

RESULT 9
BX42_DROME
ID NCBI_HUMAN STANDARD; PRT; 547 AA.
AC P39736;
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Puf specific protein Bx42.
GN BX42.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OREGON-R; TISSUE=Embryo;
RX MEDLINE=95048378; PubMed=1424996;
RA Wieland C., Mann S., von Besser H., Saunweber H.;
RT "The Drosophila nuclear protein Bx42, which is found in many puffs on
RL polytene chromosomes, is highly charged.";
RN Chromosome 101:517-525(1992).
CC -!- FUNCTION: MAY PLAY A ROLE IN CHROMATIN STRUCTURE AND FUNCTION.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DEVELOPMENTAL STAGE: TWO TRANSCRIPTS ARE DETECTED OF SIZES 1.9 AND
CC 2.2 KB. BOTH ARE DETECTED SOON AFTER FERTILIZATION AND SHOW
CC RELATIVELY CONSTANT EXPRESSION DURING THE FIRST 2/3 OF
CC EMBRYOGENESIS. IN 0-3 HR EMBRYOS, THE SMALLER TRANSCRIPT IS
CC PREDOMINANT AND THE LEVELS OF THE TWO TRANSCRIPTS ARE SOMEWHAT
CC REDUCED AT THE LATER STAGES OF DEVELOPMENT, BUT THEY ARE FOUND IN
CC APPROXIMATELY CONSTANT AMOUNTS DURING LARVAL, PUPAL AND ADULT
CC STAGES. THE SMALLER TRANSCRIPT IS SUSPECTED TO BE A MATERNAL
CC TRANSCRIPT.
CC -!- SIMILARITY: BELONGS TO THE SNW FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC EMBL; X64536; CAA45834.1; -.
DR FlyBase; FBgn0004856; Bx42.
DR InterPro; IPR004015; SKIP_SNW.
DR Pfam; PF02731; SKIP_SNW; I.
KW Nuclear protein; DNA-binding. SNW.
FT DOMAIN 177 343
FT DOMAIN 383 462 SH2-LIKE DOMAIN.
FT DOMAIN 209 236 PRO-RICH.
FT DOMAIN 125 155 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 300 400 ASP/GLU-RICH (ACIDIC).
SQ SEQUENCE 547 AA; 61156 MW; 01399EA291C9D557 CRC64;

Query Match          5.4%; Score 98; DB 1; Length 547;
Best Local Similarity 22.6%; Pred. No. 3.9;
Matches 73; Conservative 38; Mismatches 116; Indels 96; Gaps 19;

QY 44 HRDGRKQMTTEQREWLSCVGVTVWNSQYD---TNRLA-----FASFD 86
DB 234 HSPSRKTV-KEQKEWKIPPCIS-NMKNAGYTIPLDKRLAADGRGLQQVHINEKFAKVA 291
QY 87 E-----DRFKNELKNGRPSGTRAEFEGRVAKESDFEKGFORAREVASVMNRALENAH 141
DB 292 EALYIADRKARE-----AVEARSOLEKKLAQK--EKEKEDMLRMA-----QRAR 335
QY 142 DESAYLDNLKELANGNDALRNEDARSPFYSALRNTPSPFKERNGNHDPSPRMKAVIYSKH 201
DB 336 EERA-----GLRNPEAAEPSGSGATGS-EVRRN----- 363
QY 202 FWSGQDESSADRRKRYGDDPAFPAP-GTGLVDMSDNRNIPRS-----PTSPGEGFVNF 254
DB 364 ----DLRAERQERQDRNLQRAAPEKRSKLOKEREIDISEQIALGLPAKSAGNGETL 418
QY 255 DYGWFGAQTADADKTWTWTHGNHYHAP---NGSLGAMHYVESKERNWSEGY-SDFOR 307
DB 419 DQRLFNTKMGDSYGYDDEAY---NYVDKWRDSNTLGA-HIYRPSQADSDNYGGDLDA 474
QY 308 GAVYIFIP-KSWNTAPDKVKQG 329
DB 475 IVNTKRFVDPKQFSGASKEAAG 497

RESULT 10
NCBI_HUMAN
ID NCBI_HUMAN STANDARD; PRT; 1462 AA.
AC P30414;
DT 01-APR-1993 (Rel. 25, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NK-tumor recognition protein (Natural-killer cells cyclophilin-
DE related protein) (NK-TR protein).
GN NKTR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93133824; PubMed=8421688;
RA Anderson S.K., Gallinger S., Roder J., Frey J., Young H.A.,
RA Ortaldo J.R.;
RT "A cyclophilin-related protein involved in the function of natural
RT killer cells.";
RN [2]
RP REVISIONS.
RA Anderson S.K.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: COMPONENT OF A PUTATIVE TUMOR-RECOGNITION COMPLEX.
CC INVOLVED IN THE FUNCTION OF NK CELLS.
CC -!- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE VIA ITS N-TERMINUS.
```

```
CC CC 1- SIMILARITY: CONTAINS 1 CYCLOPHILIN-LIKE PEPIASE DOMAIN.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
CC EMBL; L04288; AAA35734.2; -.
CC EMBL; AF184110; AAD56402.1; -.
CC PIR; A47328; A47328.
CC HSSP; Q27450; 1A33.
CC MIN; 161565; -.
CC InterPro; IPR002130; CSA_PPIase.
CC Pfam; PF00160; pro_isomerase; 1.
CC PRINTS; PR00153; CSAPPISMRASE.
CC PROSITE; PS00170; CSA_PPIASE_1; 1.
CC PROSITE; PS00072; CSA_PPIASE_2; 1.
CC Cyclosporin; Isomerase; Rotamase; Repeat; Transmembrane.
KW DOMAIN 1 176 PPIASE, CYCLOPHILIN-TYPE.
FT DOMAIN 219 240 ARG/LYS-RICH (BASIC).
FT DOMAIN 421 457 ARG/LYS-RICH (BASIC).
FT DOMAIN 970 1010 ARG/LYS-RICH (BASIC).
FT DOMAIN 194 244 ARG/SER-RICH.
FT DOMAIN 466 574 ARG/SER-RICH.
FT DOMAIN 664 814 ARG/SER-RICH.
FT DOMAIN 1311 1348 ARG-SER TANDEM REPEAT-RICH.
SQ SEQUENCE 1462 AA; 165676 MW; D98A1147763EF527 CRC64;

Query Match 5.4%; Score 98; DB 1; Length 1462;
Best Local Similarity 20.08; Pred. No. 13;
Matches 48; Conservative 38; Mismatches 80; Indels 74; Gaps 11;

QY 9 PAEPLDRMP-----DPKRPSTGRAETV--VNNYIRKQVQYS-----H 44
Db 606 VPIPLSDPPSPRKPQKPKPSYERIQEMKAKTTHLLPIQSTYSLANIKETGSSSYH 665
QY 45 RDRGKQQTTEQREWLSCGCVTVWNSQYPTNPLAFASDFDEFRKKNELKNGRPSRGET 104
Db 666 K--REKNESDQSTYSK-----SRSSESPRSRSRSRS 699
QY 105 RAEEFEGRAVESFDEKGFQRAREVASYMRA---LENADHESAYLDNLKRELKANGDAL 161
Db 700 RS-----YSRSTRSRLASSHSRSPSRSHSRNKNYSDHSQCSRSSVTSI 747
QY 162 RNEDARSPFYSALRNPTEFKERNGNH--DPSRMKAVIYSKHFWSGGDRSSADKRYKGD 219
Db 748 SSDGGR----RAKRLRSRGKKNYSVHKHSSSEKTLHISKYV-KGRDRSCV--RKYSE 800

RESULT 11
PHOE_SALTI STANDARD; PRT; 350 AA.
AC Q56119;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Outer membrane pore protein E precursor.
GN PHOE OR STY0365
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RA Torres A., Puente J.L., Calva E.;
RC STRAIN=IMSS-1;
RL Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA Churher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Fellwell T., Hamlin N., Hague A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
RL enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
CC 1- FUNCTION: THIS IS ONE OF THE PROTEINS INDUCED WHEN CELLS ARE
CC GROWN UNDER PHOSPHATE LIMITATION. ITS PROTEIN PORE IS
CC PARTICULARLY EFFICIENT IN THE UPTAKE OF INORGANIC PHOSPHATE,
CC PHOSPHORYLATED COMPOUNDS, AND SOME OTHER NEGATIVELY CHARGED
CC SOLUTES.
CC 1- SUBUNIT: HOMOTRIMER.
CC 1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
CC 1- SIMILARITY: BELONGS TO THE OMPC/PHOE FAMILY OF PORINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
CC EMBL; X74595; CAA52672.1; -.
CC EMBL; AL627266; CAD08790.1; -.
CC HSSP; P02932; 1PHO.
CC InterPro; IPR001702; Gram_neg_porin.
CC Pfam; PF00467; Gram-ve_porins; 1.
CC PRINTS; PR00182; ECOLNEIPORIN.
CC PROSITE; PS00576; GRAM_NEG_PORIN; 1.
KW Outer membrane; Transmembrane; Porin; Signal; Complete proteome.
FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 350 OUTER MEMBRANE PORE PROTEIN E.
FT CONFLICT 63 63 E -> K (IN REF. 1).
FT CONFLICT 80 80 F -> G (IN REF. 1).
FT CONFLICT 229 229 A -> G (IN REF. 1).
FT SEQUENCE 350 AA; 38744 MW; 890F5F8C5EF0CC77 CRC64;

Query Match 5.3%; Score 96; DB 1; Length 350;
Best Local Similarity 20.5%; Pred. No. 31;
Matches 62; Conservative 33; Mismatches 125; Indels 82; Gaps 11;

QY 24 YGRAETVNNYIRKQVYSHRDRGKQQTTEQREWLSCGCVTVWNSQYPTNPLAF 83
Db 34 YGRVKAM-----HYMSDYDSKDG-----DQSYVRFQKGETQIN-----DQLTGY 73
QY 84 SFEDERFKNELKNGRPSRGETRAEFGRVAKF--SFDEKGFQRAREVASYMRALENH 141
Db 74 GRWEAEFAGNKAESDSQKTRFLAFAGLKLKDIGSFYDGNLALYDVAWTDMPFEGG 133
QY 142 DESAYLDNLKRELKANGDALRNEDARSPFYSALRNPTEFKERNGNHDPSRMKAVIYSKH 201
Db 134 DSSAQTDNEMTKRASGLATYRNTD---FFGIVDGLDLTLQYQGNEDRD-----VKKG 183
QY 202 FWSGQDRSSADKRYKGDPAEPCTGLVDMRSRNPISPRSPSGEGVNFVDYGFCA 261
Db 184 NGDGFSTVSVD---FGSD-FAVSAYTLSDTRQNLQRGT-----CD 225
QY 262 QTEADADKTVTHGNHYHAPNGSLGAMHYVESKFRNWSGYSDFDRGAYVITIPKSWT 321
Db 226 KAEA-----WATGVKYDA-----NDIYIAIFYSETRNK 253
QY 322 AP 323
Db 254 TP 255
```



```

CC      -!- SIMILARITY: BELONGS TO THE EIF3S10 FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: U14172; AAA90910.1; -
CC      EXEL: X84651; CAA59144.1; -
CC      EMBL: X17373; CAA32246.1; -
CC      PIR: S13800; S13800.
CC      MGD: MGI:95301; Eif3.
CC      InterPro: IPR000717; PCI.
CC      SMART: SM00088; PINT: 1.
CC      Initiation factor; protein biosynthesis. Repeat: Phosphorylation.
CC      DOMAIN 924 1133
CC      [GP]-[PS]-[RW]-R-[GN]-[AM].
CC      21 X 10 AA TANDEM REPEAT OF D-[DE]-D-R-
CC      [GK]AREKEREILQEHQIKKTVRERLEQIKKTE ->
CC      PRKGAREGTNPSTRANQENGSGAVRADQEDR (IN
CC      REF. 2 AND 3).
CC      EL -> DY (IN REF. 2 AND 3).
CC      Q -> H (IN REF. 2 AND 3).
CC      A -> V (IN REF. 2 AND 3).
CC      RHR -> SIVA (IN REF. 3).
CC      E -> D (IN REF. 2).
CC      SEQUENCE 1344 AA; 161949 MW; F4CAE2169F577712 CRC64;
CC      -----
CC      Query Match 5.3%; Score 96; DB 1; Length 1344;
CC      Best Local Similarity 18.9%; Pred. No. 17;
CC      Matches 53; Conservative 39; Mismatches 103; Indels 86; Gaps 11;
CC      -----
CC      QY 3 DDRVTPAEPLDRMPDYPYRSYGAETVNNYIRK-WQQVYSHRGR-----KQQTTEEQ 56
CC      Db 1115 DDRV--PRRGDARPGWPR-----FVKPGWREKEKAREESWGPPRESRPSSE 1161
CC      QY 57 REWLSYGVGVTVWNSQYPTNRLAFASDFEDRFKFKELKNGRPSRGSTRAEFEGRVAKES 116
CC      Db 1162 REW-----DRDKEKDRDNQDRENDKDLDRDRERD 1193
CC      QY 117 FDEKGFQRAEVASVNNRALENHDSAYLDNLKELANGNDALRNEDARSFFYSALRN 176
CC      Db 1194 GDREDFRFRPDEGGWRGP---AESSSWRDSRRDRDRDRDRDRDRDRDRDRDRDRDR 1243
CC      QY 177 TPSTFKERNGNHDPSPRMKAVIYKSH---FWSGDRSSSADKRYGDPDAFRAPGTGLV 232
CC      Db 1244 LRDLDRDLDRDLDRDRGFLRSEREREASSWRRTD-DRKDRTEERDPPRRVPPPA---- 1298
CC      QY 233 DMSRDRNIPRSPSPGEGFVNFYWGFGAQTEADAKTYW 272
CC      Db 1299 -LSRDRERER-----BREGKEKASW 1318
CC      -----
CC      RESULT 14
CC      MAPB_HUMAN STANDARD; PRT; 2468 AA.
CC      AC P46821;
CC      DT 01-NOV-1995 (Rel. 32, Created)
CC      DT 01-NOV-1995 (Rel. 32, Last sequence update)
CC      DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC      DE Microtubule-associated protein 1B (MAP 1B) [Contains: MAP1 light chain
CC      LCL]
CC      GN MAP1B.
CC      OS Homo sapiens (Human).
CC      OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC      OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC      CX NCBI_TaxID=9606;
CC      RN [1]
CC      RP SEQUENCE FROM N.A.

```

```

RC      TISSUE=Fetal brain;
RC      MEDLINE=95104835; PubMed=7806212;
RA      Lien L.L., Feener C., Fischbach N., Kunkel L.M.;
RT      "Cloning of human microtubule-associated protein 1B and the
RL      identification of a related gene on chromosome 15.";
RL      Genomics 22:273-280(1994).
CC      -!- FUNCTION: THE FUNCTION OF BRAIN MAPS IS ESSENTIALLY UNKNOWN.
CC      THAT ACCOMPANY MAP1B MAY PLAY A ROLE IN THE CYTOSKELETAL CHANGES
CC      PHOSPHORYLATED MAP1B EXTENSION. POSSIBLY MAP1B BINDS TO AT LEAST
CC      TWO TUBULIN SUBUNITS IN THE POLYMER, AND THIS BRIDGING OF SUBUNITS
CC      MIGHT BE INVOLVED IN NUCLEATING MICROTUBULE POLYMERIZATION AND IN
CC      STABILIZING MICROTUBULES.
CC      -!- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LCL1, LCL2 AND LCL3, CAN ASSOCIATE
CC      WITH MAP1A AND MAP1B PROTEINS.
CC      -!- DOMAIN: Has a highly basic region with many copies of the sequence
CC      KKEE and KKEI/V, repeated but not at fixed intervals, which is
CC      responsible for the binding of MAP1B to microtubules.
CC      -!- PTM: LCL1 is coexpressed with MAP1B. It is a polypeptide generated
CC      from MAP1B by proteolytic processing. It is free to associate with
CC      both MAP1A and MAP1B. It interacts with the amino-terminal region
CC      of MAP1B (BY SIMILARITY).
CC      -!- SIMILARITY: TO MAP1A.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: L06237; AAA18904.1; -
CC      MIM: 157129; -
CC      InterPro: IPR000102; MAP1B_neuraxin.
CC      Pfam: PF00414; MAP1B_neuraxin; 10.
CC      PROSITE: PS00230; MAP1B_NEURAXIN; 6.
CC      Microtubules; Repeat; Phosphorylation.
CC      CHAIN ? 2468 MAP1 LIGHT CHAIN LCL1.
CC      FT REPEAT 1878 1894 MAP1B 1.
CC      FT REPEAT 1895 1911 MAP1B 2.
CC      FT REPEAT 1912 1928 MAP1B 3.
CC      FT REPEAT 1929 1945 MAP1B 4.
CC      FT REPEAT 1946 1962 MAP1B 5.
CC      FT REPEAT 1963 1979 MAP1B 6.
CC      FT REPEAT 1997 2013 MAP1B 7.
CC      FT REPEAT 2014 2030 MAP1B 8.
CC      FT REPEAT 2031 2047 MAP1B 9.
CC      FT REPEAT 2048 2064 MAP1B 10.
CC      FT DOMAIN 589 790 LYS-RICH (HIGHLY BASIC, CONTAINS MANY
CC      KKEE AND KKEI/V REPEATS).
CC      SEQUENCE 2468 AA; 270618 MW; 540839C8DF09D461 CRC64;
CC      -----
CC      Query Match 5.3%; Score 96; DB 1; Length 2468;
CC      Best Local Similarity 21.9%; Pred. No. 37;
CC      Matches 68; Conservative 39; Mismatches 110; Indels 94; Gaps 14;
CC      -----
CC      QY 46 DGRKQOMTEEQREWLSYG---CVGVTVWNSQYPTNRLAFASDFEDRFKFKELKNGRPRS 101
CC      Db 941 DYEEKAETEEAEPEDEEGEEHVCSA---SKHSPTDEESAKADAYIRE-KRESVAS 995
CC      QY 102 GETRAE-----FEGRVAKESFDEKGFQARE-----VASVMNRALE- 138
CC      Db 996 GDDRAEEDMDAELEKGEAEQSEEADEEDKADAREEVEPEKMEADYVMVAVDKAAEA 1055
CC      QY 139 -NAHDESAYLDNLKELANGNDALRNEDARSFFYSALRNTFSFKERNGNHDPSPMKAVI 197
CC      Db 1056 GAEGSOYGLTPTTKOLG-----AQSP-----GREPASSIHDETLPGG-- 1093
CC      QY 198 YSKHFWSGODRSSADKRYGDPDAFRPAPG-----TGLVDMRSDRNIIPRSP 245
CC      Db 1094 -----SESEATASDENREDQPEEFTATSGVTQSTQSTIISSEPTMDENSTPRDWSDET 1147

```

Search completed: September 27, 2002, 12:39:57
Job time: 105 sec